Final Report for Period: 07/2009 - 06/2010
Submitted on: 07/06/2010
Principal Investigator: Kvam, Paul H.
Award ID: 0700131
Organization: GA Tech Res Corp - GIT
Submitted By:

Title:
Collaborative Research: Modeling Reliability for Scale-Driven Degradation and Spatial Defects

Project Participants

Senior Personnel

Name: Kvam, Paul
Worked for more than 160 Hours: Yes
Contribution to Project:

Post-doc

Name: Quan, Sun
Worked for more than 160 Hours: No
Contribution to Project:
Professor Quan was a visiting professor who joined PI's research team in 2009. He was funded through University of Defense Technology, Changsha, China

Graduate Student

Name: Thelen, Michael
Worked for more than 160 Hours: Yes
Contribution to Project:
New graduate student at Georgia Tech who is learning basics of statistics in nanomanufacturing. His participation is limited to reading and programming.

Undergraduate Student

Technician, Programmer

Other Participant

Research Experience for Undergraduates

Organizational Partners

Other Collaborators or Contacts

On the other PIs listed on the collaborative proposal (they file separate reports)

I collaborated with Suk Joo Bae of Hanyang University (Seoul, Korea) about research related to this grant.

I collaborated with Sun Quan of University of Defense Technology, Changsha, China

I am in communication with Shiren Wang, Texas Tech University,
with regard to research related to this grant.

Activities and Findings

Research and Education Activities:

1. Studying length bias in the measurements of carbon nanotubes. General result will help to calibrate software that is currently marketed to nanotechnology labs across the world.

2. Modeling hot electron degradation in nano-scaled MOSFET devices using parametric change-point models and nonparametric errors. This is in collaboration with Suk Joo Bae.

3. Modeling Rubidium lamp using the same multiple-cause degradation modeling framework proposed in NSF grant (originally proposed for MOSFET devises). This is in collaboration with Sun Quan

4. Working with Yue Kuo's students at his Thin Film Nano & Microelectronics Research Laboratory in which the activities include new materials, novel processes, and advanced devices. My focus is on reliability statistics at various stages of device development.

5. Working with my student on adapting nonparametric statistical process control techniques toward lab development of nano devices. Because nanomanufacturing does not have a stable, traditional process control regimen, we focus on summarizing multivariate output using distribution free methods, also considering Bayesian methods in order to combine disparate experimental information.

6. Presentations: Presented findings of length bias in the measurements of carbon nanotubes at Naval Postgraduate School (1/08), Texas A&M (3/08), Auburn University (10/08)

Findings: (See PDF version submitted by PI at the end of the report)

1. For carbon nanotube length bias: To measure carbon nanotube lengths, atomic force microscopy and special software are used to identify and measure nanotubes on a square grid. Current practice does not include nanotubes that cross the grid, and as a result, the sample is length biased. The selection bias model can be demonstrated through Buffon's Needle Problem, which was extended to general curves that more realistically represent the shape of nanotubes observed on a grid. In this paper, the nonparametric maximum likelihood estimator is constructed for the length distribution of the nanotubes, and the consequences of the length bias are examined. Probability plots reveal that the corrected length distribution estimate provides a better fit to the Weibull distribution than the original selection-biased observations, thus reinforcing a previous claim about the underlying distribution of synthesized nanotube lengths.

2. Statistical Models for Hot Electron Degradation in Nano-Scaled MOSFET Devices: In a MOS structure, the generation of hot carrier interface states is a critical feature of the device's reliability. On the nano-scale, there are problems with degradation in transconductance, shift in threshold voltage, and decrease in drain current capability. Quantum mechanics has been used to relate this decrease to degradation and device failure. Although the lifetime and degradation of a device are typically used to characterize its reliability, in this paper we model the distribution of hot-electron activation energies, which has appeal because it exhibits two-point discrete mixture of logistic distributions. The logistic mixture presents computational problems that are addressed in simulation.

For modeling Multi-Cause Degradation for Rubidium Lamps: At the core of satellite rubidium standard clocks is the rubidium lamp, which is a critical piece of equipment in a satellite navigation system. There are many challenges in understanding and improving the reliability of the rubidium lamp, including the extensive lifetime requirement and the dearth of samples available for destructive life tests. Experimenters rely on degradation experiments to assess the lifetime distribution

of highly reliable products that seem unlikely to fail
under the normal stress conditions, because degradation data can provide extra information about product reliability. Based on recent research on the rubidium lamp, this article presents a
multi-causes degradation path model, including its application background, model description, modeling method, and parameter estimation method. Using available data from degradation tests,
we construct point estimates and interval estimates for rubidium lamp lifetimes using regression techniques.

Training and Development:
As part of his final report, Ph.D. student Michael Thelen gave a 45 minute presentation on reliability and statistical process control in nanomanufacturing. It was his first peer-attended talk as a graduate student.

Sun Quan presented our findings to the Georgia Tech School of Industrial and Systems Engineering Statistics Seminar in Fall, 2009. It was his first English presentation outsided of China.

Outreach Activities:

Journal Publications

Bae, Kim, Kuo, Kvam, "Statistical Models for Hot Electron Degradation in Nano-Scaled MOSFET Devices", IEEE Transactions on Reliability, p. 392, vol. 57, (2008). Published,

Kvam, "Length Bias in the Measurements of Carbon Nanotubes", Technometrics, p. , vol. , (2009). Accepted,

Kvam, P. H. and Quan, S., "Multi-Causes Degradation Path model: A case study on rubidium lamp Degradation", Quality and Reliability Engineering International, p. , vol. , (2010). Submitted,

Books or Other One-time Publications

Web/Internet Site

Other Specific Products

Contributions

Contributions within Discipline:
Technometrics paper showed sampling defects of Simagas software in measuring Nanotube lengths.

Quality and Reliability Engineering International paper presents first degradation model of this kind applied toward rubidium clock reliability in sattelites.

Contributions to Other Disciplines:
By showing problem with current nanotube measurement software, many past results will need to be recalibrated, which could change scientific understanding about quality and characterisitcs of CNTs produced in a lab.

Contributions to Human Resource Development:

Contributions to Resources for Research and Education:

Contributions Beyond Science and Engineering:
Technometrics paper: By finding flaw in Simigas software, we should inspire corrective action and possibly help competative technologies

aspire to better CNT measuring frameworks for laboratory work.

Quality and Reliability Engineering International: Current paper is spurring on three different research projects at UDT in Changsha based on a multi-cause degradation path model.

Conference Proceedings

Categories for which nothing is reported:

Organizational Partners

Activities and Findings: Any Outreach Activities

Any Book

Any Web/Internet Site

Any Product

Contributions: To Any Human Resource Development

Contributions: To Any Resources for Research and Education

Any Conference

Length Bias in the Measurements of Carbon Nanotubes

Paul KVAM

H. Milton Stewart School of Industrial Engineering
Georgia Institute of Technology
Atlanta, GA

To measure carbon nanotube lengths, atomic force microscopy and special software are used to identify and measure nanotubes on a square grid. Current practice does not include nanotubes that cross the grid, and, as a result, the sample is length-biased. The selection bias model can be demonstrated through Buffon's needle problem, extended to general curves that more realistically represent the shape of nanotubes observed on a grid. In this article, the nonparametric maximum likelihood estimator is constructed for the length distribution of the nanotubes, and the consequences of the length bias are examined. Probability plots reveal that the corrected length distribution estimate provides a better fit to the Weibull distribution than the original selection-biased observations, thus reinforcing a previous claim about the underlying distribution of synthesized nanotube lengths.

KEY WORDS: Buffon's needle problem; Censoring; Goodness of fit; Nonparametric; Probability plot; Sampling; Selection bias; Weibull; Weighted distributions.

1. INTRODUCTION

Carbon nanotubes are rolled-up nanoscale sheets of graphitic carbon used to enhance materials with tensile strength, thermal conductivity, and electronic conductivity. In terms of tensile strength alone, they can be stiffer and stronger than any other microscaled fiber. Nanotubes and similar nanosized devices have garnered great attention in the industrial and academic world, even though statistical research in nanomanufacturing is still in its early stages. In contrast to laboratory research, commercial applications have developed at a slow pace, due primarily to the enormous production costs of high-quality nanotubes.

A major challenge in the production of nanotubes is controlling and analyzing the important qualities, such as length, diameter, and strength. Nanotube length can be measured using atomic force microscopy (AFM) and special software developed by SIMAGIS, which processes three-dimensional images of bulk nanotube material samples. The software identifies and classifies nanotubes and then analyzes nanotube lengths while recognizing the intersections, contacts, and three dimensions of curvature of the tubes. The solution features also include the ability to filter objects according to selected length and height. In this article we focus on length measurements. Figure 1 presents an example of an AFM profile image of carbon nanotubes in a 5 μm × 5 μm square interval. The average size of these nanotubes is nearly .5 μm, or 500 nm. This recent technological innovation gives experimenters the ability to ascertain distributional properties of carbon nanotubes beyond the simple mean lengths, as characterized by Ziegler et al. (2005). Important mechanical properties that are attainable include tensile strength, elastic modulus, and fracture toughness.

Wang, Liang, Wang, and Zhang (2006) presented a statistical characterization of nanotube length using an AFM sample and the SIMAGIS software. The set of length measurements for 651 detected nanotubes is summarized in a histogram in Figure 2. Like the image in Figure 1, the sampling area is a square grid with sides measuring 5 μm. Although it has been verified in smaller samples that the software accurately finds and measures nanotubes smaller than 10 nanometers, the software's count does not include nanotubes that cross the boundary of the square, and thus the data are length-censored. Although the censoring mechanism can be easily included in the statistical analysis, this sampling method introduces size bias into the analysis as well.

2. LENGTH BIAS

For a positive random variable X with density function $f(x)$, if the random variable is observed with a probability proportional to its size $w(x)$, then the distribution of the random sample obtained with this bias has density

$$f_w(x) = \frac{w(x)f(x)}{\int w(x)f(x)\,dx}.$$

This kind of weighted bias in sampling was first discussed by Rao (1965), with a particular interest in length bias, where $w(x) = x$. This holds for discrete distributions as well, and we use $dF(x)$ to denote the more general class of distributions.

Length bias is a common phenomenon in survival studies in which prevalent cases are determined through cross-sectional study; for example, a test tends to detect more slow-growing cancers that take longer to become symptomatic compared with aggressive, fast-growing cancers. Cancers that grow slowly are easier to detect because they have a longer presymptomatic period when they are detectable, and thus show up with more relative frequency in such cancer studies. Horvath (1985) and Asgharian (2002) have presented a more detailed discussion on sampling and analysis problems induced by length bias.

TECHNOMETRICS, ???? 0, VOL. 0, NO. 0
DOI 10.1198/004017008000000442

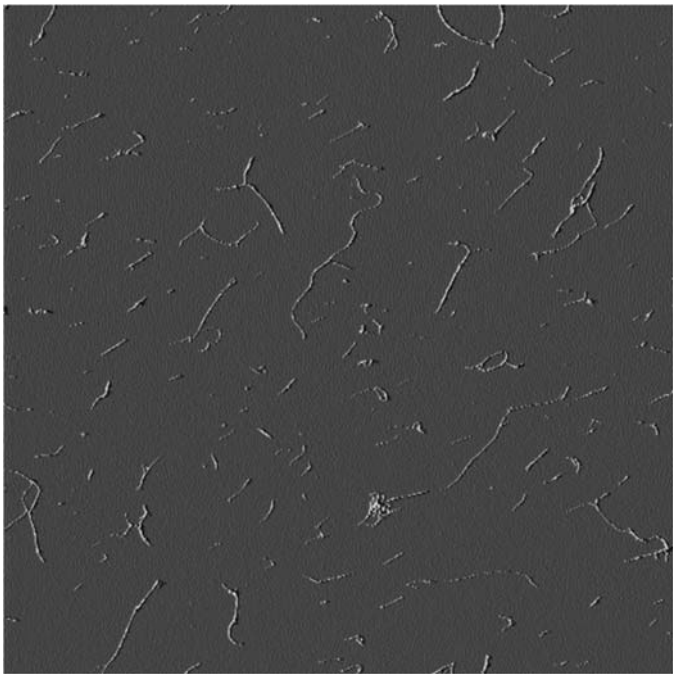


Figure 1. AFM image of carbon nanotubes.

2.1 Buffon's Needle Problem

Buffon's famous needle problem can be used to illustrate the length bias that occurs in this kind of two-dimensional sample. In 1733, Georges-Louis Buffon, a French naturalist and mathematician, posed the following problem: Given a needle of length r and a grid of parallel lines spaced by an equal distance a, what is the probability that a randomly tossed needle will cross one of these lines? The problem is solved with rudimentary geometry based on an assumption that the center of the needle and its angle with respect to the grid are uniformly distributed. This example has been used as a classic probability problem and as an experimental way of obtaining an estimate of the constant π (see, e.g., Perlman and Wichura 1975). In this article we focus on the case in which $r \leq a$, where the censoring probability is $2r/(a\pi)$.

By 1812, Pierre-Simon Laplace considered a second grid of parallel lines constructed at right angles to the first set of parallel lines and with the same distance a, and found that the crossing probability doubled; that is, the probability of a line crossing on a square boundary is $4r/(a\pi)$. A geometrical proof was offered by Mantel (1952).

Gnedenko (1962) provided a more general extension of the crossing probability to multisided convex polygons under the constraint that the diameter is $<a$. Ramely (1969) reframed this extension in terms of arbitrary curves; this is sometimes referred to as "Buffon's noodle problem." With polygons and crooked lines, the crossing frequency changes, but the number of expected crossings remains constant. To illustrate this, think of a needle bent in half. The probability the folded-over needle crosses a grid line is halved, but the expected number of crossings stays the same, because the the needle will necessarily cross the the grid line twice.

To represent a nanotube of length r, we could treat it as a straight needle with crossing probability of $4r/(a\pi)$ or as a crooked, curving tube that connects to itself, in which case the crossing probability is reduced to $2r/(a\pi)$. As shown in Figure 2, the nanotubes exhibit only a modest curvature, much more like a straight needle than a connected polygon.

To illustrate the effect of curvature in a simple way, we assume that any nanotube consists of two segments of equal length. Figure 3 shows the nanotubes with segment angles of $\theta = \{0, \pi/8, \pi/4\}$. If a nanotube consists of two segments of length $r/2$, then the sum of the three side lengths of the inscribed triangle is $r(1 + \cos(\theta))$. Accordingly, the crossing probability is reduced to

$$\frac{4r}{a\pi}\lambda, \tag{1}$$

where $\lambda = (\cos(\theta) + 1)/2$ represents the shrinkage in the crossing probability due to the bending of the nanotube. For $\theta = \{0, \pi/8, \pi/4\}$, we have $\lambda = \{1, .9619, .8536\}$. It might be sensible to treat λ as a random effect, reflecting the apparent heterogeneity of shapes observed in the sample of synthesized

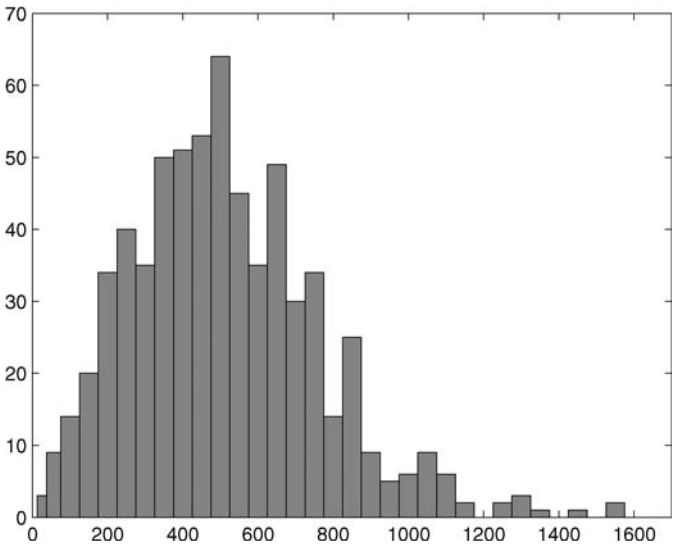


Figure 2. Histogram of dispersed SWNT lengths.

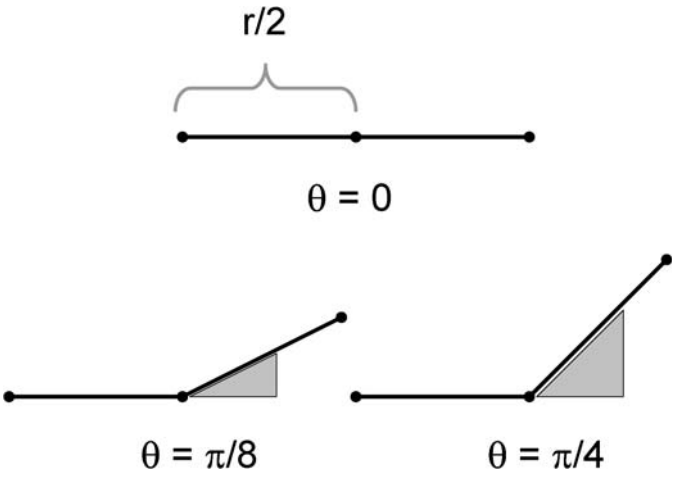


Figure 3. Nanotubes represented as bisected needles with angles $\theta = \{0, \pi/8, \pi/4\}$.

nanotubes. In this article we focus on a fixed, known λ, emphasizing the case of $\lambda = 1$ (straight nanotubes). The effects of curvature and randomness in the shape properties are discussed further in Section 4.

2.2 Censored Nanotubes

Obviously, a longer nanotube is more likely to be censored (and thus not measured) than a smaller one. Consequently, a sample that ignores the censored lines is biased according to the nanotube's length. Suppose that the carbon nanotube lengths have cumulative distribution function $F_L(t)$. We assume $F_L(t_0) = 1$ for some $t_0 < a$, although this constraint is not necessary for demonstrating the selection bias. Let $\mu_L = \int t\, dF_L(t)$. Based on the approximation in (1), the overall probability of censoring is

$$p = \int \frac{4\lambda t}{a\pi}\, dF_L(t) = \frac{4\lambda\mu_L}{a\pi}. \tag{2}$$

From the software output, we observe a sample of n nanotube measurements that has empirical distribution function $\tilde{F}_n(t)$. Let F_0 and F_1 be the distribution functions of the censored observations (nanotubes that cross the boundary) and uncensored observations.

From the mixture

$$F_L(t) = pF_0(t) + (1-p)F_1(t), \tag{3}$$

$\tilde{F}_n(t)$ is actually the nonparametric maximum likelihood estimator (NPMLE) for F_1, not F_L. If the censoring probability p remains small, then the bias from estimating $F_L(t)$ with $\tilde{F}_n(t)$ might be minuscule. For problems in which the nanotube length represents a more significant proportion of a, however, the bias can be sizeable, as we show later. From (3), the length bias $[(w(x) = x)]$ affects both the censored population (F_0) and the observed population (F_1) as

$$\begin{aligned} dF_0(t) &= \frac{t}{\mu_L}\, dF_L(t) \qquad \text{and} \\ dF_1(t) &= \frac{a\pi - 4\lambda t}{a\pi - 4\lambda\mu_L}\, dF_L(t), \end{aligned} \tag{4}$$

where $0 \le t \le a$. Whereas the censored observations are length-biased, the observations lying strictly within the square also are biased, but in an opposite manner.

In the remainder of the article we restrict our consideration to distributions for which both F_0 and F_1 are proper distributions functions. From (4), we require that $F_L(t_0) = 1$ for some $t_0 < a\pi/4$.

2.3 Example

A simple example illustrates how length bias can adversely affect the sample outcome if one does not compensate for it. Suppose that the underlying distribution for (straight) nanotube lengths is Uniform$(0, 2\mu)$, observed on a unit grid ($a = 1$). From $dF_1(t)$ in (4), the density function of the uncensored nanotubes is

$$f_1(x) = \left(\frac{1}{2\mu}\right)\frac{\pi - 4x}{\pi - 4\mu}, \qquad 0 \le x \le 2\mu, \tag{5}$$

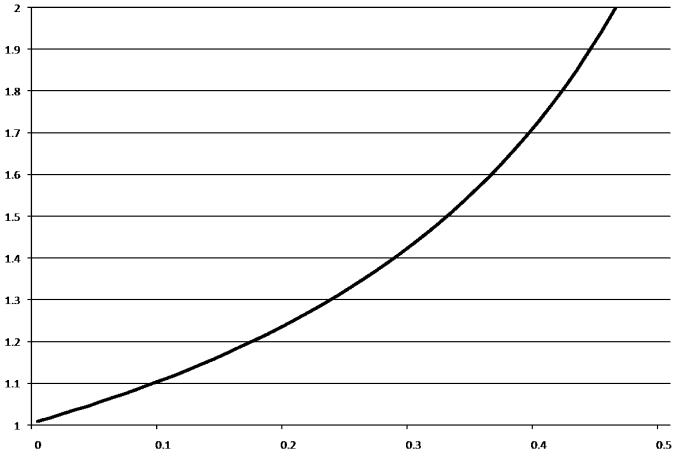


Figure 4. Ratio of μ_L over the mean of $F_1(t)$ (μ_1) for nanotube mean lengths up to 50% of grid size.

where μ is the mean nanotube length. In this case the mean for $F_1(t)$ is

$$\mu_1 = \mu\left(\frac{3\pi - 4\mu}{3\pi - 12\mu}\right).$$

Figure 4 shows the ratio mean for $F_L(t)$ over the mean of $F_1(t)$.

Once the mean nanotube length increases to 20% of the length of the grid, the increase in μ_1 will quickly get out of hand. We further investigate the effect of length biasing for the nanotube data in the following section, where we consider nonparametric estimation of the underlying length distribution.

3. ESTIMATING THE LENGTH DISTRIBUTION

Although the bias for $dF_1(t)$ in (4) is not the commonly observed form of length bias, it nonetheless presents a straightforward selection-bias estimation problem that can be solved directly. If our observed sample is denoted by $x_1, \ldots, x_n$, then, from Vardi (1985), the biased distribution can be expressed as some distribution function $G(t)$ such that

$$G(t) = \frac{\int_0^t w(u)\, dF_L(u)}{\int_0^\infty w(u)\, dF_L(u)},$$

where the weight function representing the bias is $w(x) = a\pi - 4\lambda x$. In this case $G(t)$ is identifiable only for values of $t < a\pi/(4\lambda)$.

As a result, it can be shown, from theorems 1 and 2 of Vardi (1985), that the NPMLE of F_L is

$$\hat{F}_L(t) = \frac{\sum_{i=1}^n I(x_i \le t)(a\pi - 4\lambda x_i)^{-1}}{\sum_{i=1}^n (a\pi - 4\lambda x_i)^{-1}}. \tag{6}$$

3.1 Properties of $\hat{F}_L$

Borrowing the method of proof from theorem 3.2 of Vardi (1982), we can show the following result. We assume the coefficient representing nanotube curvature λ is fixed and known.

Theorem 1. Assume that $F(t)$ is continuous for values $0 < t < a\pi/(4\lambda)$ and has bounded density with $F(0) = 0$ and $F(a\pi/(4\lambda)) = 1$. Then

$$\int_0^{a\pi/(4\lambda)} (a\pi - 4\lambda x)\, d\hat{F}_L(x) \longrightarrow \int_0^{a\pi/(4\lambda)} (a\pi - 4\lambda x)\, dF_L(x) \quad \text{with probability 1,}$$

and

$$\sqrt{n}(\hat{F}_L(t) - F_L(t))$$

converges in distribution to a pinned Gaussian process with mean 0 and covariance function $C(u, v)$, where, for $u \le v \le r$ and $\psi(t) = \int_0^t (a\pi - 4\lambda s)^{-1}\, dF_L(s)$,

$$C(u, v) = (a\pi - 4\lambda\mu_L) \times \big(\psi(u)(1 - F_L(v)) + F_L(u)(\psi(r)F_L(v) - \psi(v))\big).$$

3.2 Nanotube Data

In the analysis of the data of Wang et al. (2006), we use $\lambda = 1$ and refrain from making more elaborate assumptions about the apparent nanotube curvature displayed in the AFM image. Figure 5 shows the estimated survival function along with an approximate 95% confidence interval, based on the nonparametric percentile bootstrap procedure (with 1,000 bootstrap samples) described in chapter 15.2 of Kvam and Vidakovic (2007). Because of the large sample size, the coverage probability is not significantly improved using bias correction.

To see how the corrected nonparametric estimator compares with the uncorrected one, we can visually compare their density estimates. Figure 6 shows a smoothed version of the density associated with $\hat{F}_L(t)$ (dotted line) alongside the biased estimator constructed from the unadjusted empirical density (solid line). The estimated mean and standard deviation using the standard estimator (530.5 nm, 261.4 nm) are not remarkably different from the mean and standard deviation of the data (510.4 nm, 253.2 nm). In this case the mean nanotube length is on the order of 10 times smaller than the length of the side of the square from which they are sampled.

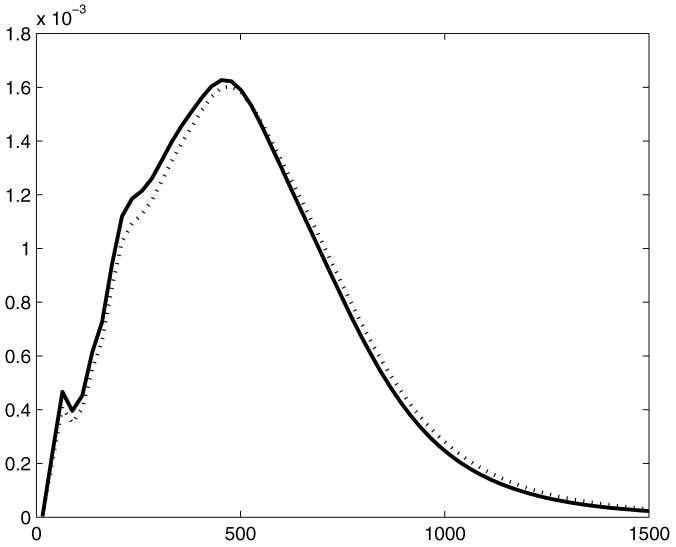


Figure 6. Empirical density (——) and adjusted density estimator (· · ·) for carbon nanotube lengths.

Obviously, if the nanotube lengths are more in proportion to the side of the square, then the length bias can be much more dramatic. Figure 7 shows the original empirical density and estimators adjusted for two smaller (hypothetical) AFM image squares. The squares have sizes measuring 2,500 and 2,000 nm (roughly between four and five times larger than the mean nanotube size). The estimated densities show how the NPMLE adjusts for censoring of more of the larger nanotubes as the window size is reduced. For the 2,500 nm^2 window, this is reflected in mass being transferred to the right, illustrating how substantial bias can be created by the missing censored observations.

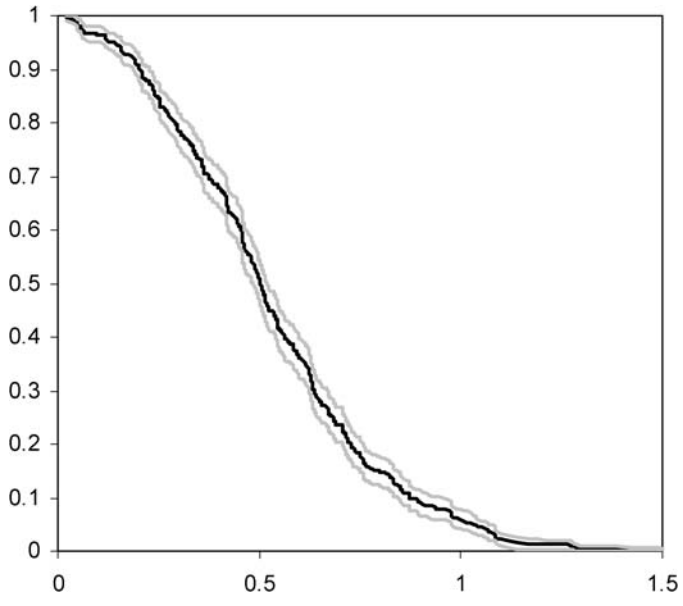


Figure 5. Nonparametric survival function (black line) and 95% confidence interval (gray lines) for carbon nanotube lengths.

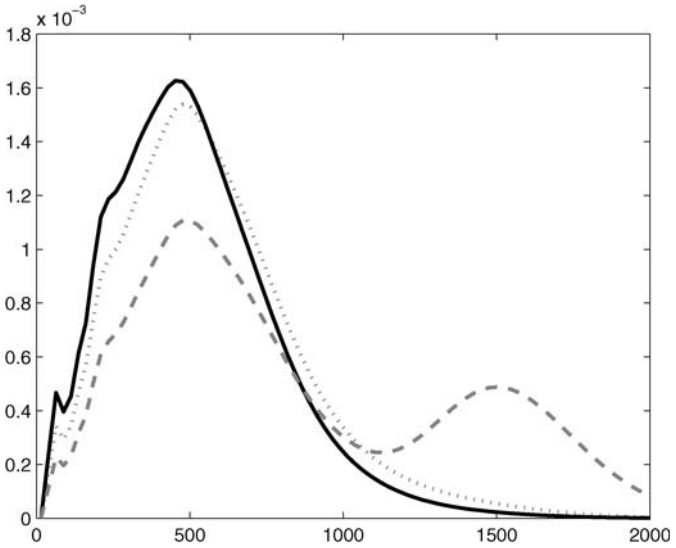


Figure 7. Empirical density (——) and adjusted density estimators based on different AFM image sizes. Original size is 5,000 nm. Adjusted sizes are 2,500 nm (· · ·) and 2,000 nm (- - -).

For the 2,000 nm^2, the NPMLE has generated an extra mode for probability mass representing the lost observations due to censoring.

3.3 Effect of Shape

With sound knowledge about the curvature of nanotubes in the AFM image, models based on $\lambda < 1$ suggest that the estimation of the length distribution can be improved. For the carbon nanotube data, however, changing the model by assigning $\lambda = .9$ will not change the NPMLE noticeably; for example, the estimated mean changes by only .68%. Estimating the sample curvature has not been addressed in the nanomanufacturing literature, even though it poses an intriguing problem for future research. More importance has been placed on tube lengths, diameters, and clusters than on how the nanotubes curve and bend.

Along with unknown curvature, heterogeneity within synthesized batches of carbon nanotubes has been carefully researched (see, e.g., Lu and Bhattacharya 2005). Again, the randomness of the nanotube curvature has not received special attention beyond how the tubes will overlay one another, which naturally affects the electrical properties of nanodevices. Although the metric implied in Figure 3 is overly simplistic in many regards, here it adequately serves the purpose of illustrating the effects of shape variability on the censoring, length-biasing, and nonparametric estimation of nanotube length.

In this example the carbon nanotubes are represented as bisected needles, and the angle of bend θ is governed by a uniform distribution, $\theta \sim \mathrm{U}(0, \pi/4)$. The resulting distribution of λ (related to the arc since distribution) percolates uncertainty down through $\hat{F}_L$. In this example the length distribution is $\mathrm{U}(0, 2\mu)$, and F_1 is represented as a mixture with mixing density $g_\lambda(w) = 4\pi^{-1}(1 - w^2)^{-1/2}$, corresponding to $\lambda = (\cos(\theta) + 1)/2$. Through numerical integration, we can compare the variation in noncensored observations from the mixture. If $\mu \leq .25$, then the standard deviation of the mixture sample is within 1% of the sample with fixed $\lambda = \pi/8$, suggesting that mild differences in nanotube curvature will not significantly affect the distribution of noncensored observations.

4. DISCUSSION OF RESULTS

The problem of locating and measuring nanosized matter has been greatly improved in recent years due to improving AFM technology and recognition software. But if the sampling mechanism is unable to include observations crossing the boundary of the observation grid, then the resulting length bias in the data set must be dealt with. If the bias is not addressed, then the length distribution for nanotubes will be underestimated (along with the variability). In the example of Wang et al. (2006), the consequences of the bias are relatively benign; however, other examinations of synthesized carbon nanotubes have included smaller regions relative to the average size of the nanotubes. The length of synthesized nanotubes are increasing as the technology improves; Burke, Li, and Yu (2006) describe carbon nanotubes of .4 cm in length that have been developed in research labs. Our analysis shows that the bias will have a substantial effect on the estimator of the length distribution if the mean nanotube length is as high as 20% of the size of the side of the square from which they are sampled.

Note that if censored observations were measured, then the analysis would be straightforward using the Kaplan–Meier product limit estimator. In some cases a nanotube located in the corner of a square can cross both grid boundaries, but in the nature of sampling nanotube lengths, the (right) censoring is the same.

Wang et al. (2006), set out to prove that the underlying length data can be characterized by the Weibull distribution. Indeed, the Weibull model has served as a precedent for similar distributional properties of fibrous materials. Based on the model for fiber composites of Fukuda and Kawata (1974) and Jayaraman and Kortschot (1996), Wang et al. (2006) showed that the epoxy composite stiffness measurements are consistent with Weibull-distributed nanotube lengths. Although their probability plot of the original data in Figure 8 is not entirely convincing, one may conclude that the Weibull presents a "reasonable" distribution fit to the sample. Knowing that the data are length-biased, we can reconsider this goodness of fit for the NPMLE, $\hat{F}_L$. Comparing the original Weibull probability plot in Figure 8 to the corrected one in Figure 9, we see that the Weibull fit actually has improved slightly. This is confirmed by the decrease in the computed Kolmogorov–Smirnov test statistics; that is, the corrected estimator for the length-biased data offers stronger evidence for the conjecture of Wang et al. (2006).

ACKNOWLEDGMENTS

This research was supported by National Science Foundation grant CMMI-0700131. The author thanks the associate editor and referees for their helpful comments that led to several improvements to the article.

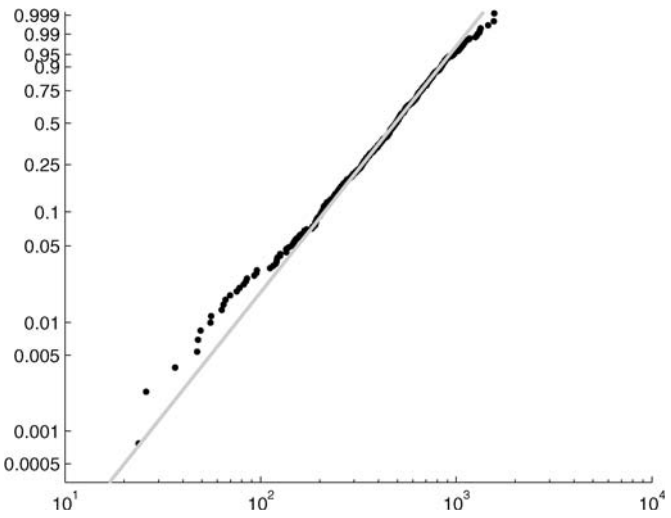


Figure 8. Original Weibull probability plot of the length-biased data for carbon nanotube lengths.

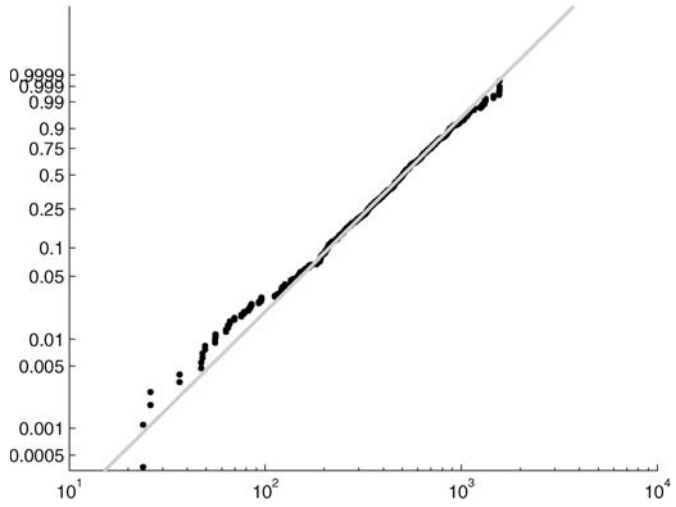


Figure 9. Weibull probability plot of the adjusted estimator of the carbon nanotube length distribution.

[Received ???. Revised ???.]

REFERENCES

Asgharian, A., M'Lan, C. E., and Wolfson, D. B. (2002), "Length-Biased Sampling With Right Censoring: An Unconditional Approach," *Journal of the American Statistical Association*, 97, 201–209.

Burke, P. J., Li, S., and Yu, J. (2006), "Quantitative Theory of Nanowire and Nanotube Antenna Performance," *IEEE Transactions on Nanotechnology*, 5, 314–334.

Fukuda, H., and Kawata, K. (1974), "On Young's Modulus of Short Fiber Composites," *Fibre Science and Technology*, 7, 207–212.

Gnedenko, B. V. (1962), *The Theory of Probability*, Moscow: MIS.

Horvath, L. (1985), "Estimation From a Length-Biased Distribution," *Statistics and Decisions*, 3, 91–113.

Jayaraman, K., and Kortschot, M. T. (1996), "Correction to the Fukuda–Kawata Young Modulus Theory and the Fukuda–Chao Strength Theory for Short Fiber–Reinforced Composite Materials," *Journal of Materials Science*, 31, 2059–2064.

Kvam, P. H., and Vidakovic, B. (2007), *Nonparametric Statistics With Applications to Science and Engineering*, Hoboken, NJ: Wiley.

Lu, Q., and Bhattacharya, B. (2005), "The Role of Atomistic Simulations in Probing the Small-Scale Aspects of Fracture: A Case Study on a Single-Walled Carbon Nanotube," *Engineering Fracture Mechanics*, 72, 2037–2071.

Mantel, N. (1952), "An Extension of the Buffon Needle Problem," *Annals of Mathematical Statistics*, 24, 674–677.

Perlman, M. D., and Wichura, M. J. (1975), "Sharpening Buffon's Needle," *The American Statistician*, 29, 157–163.

Ramaley, J. F. (1969), "Buffon's Noodle Problem," *American Mathematical Monthly*, 76, 916–918.

Rao, C. R. (1965), "On Discrete Distributions Arising Out of Methods of Ascertainment," in *Classical and Contagious Discrete Distributions*, ed. G. P. Patil, Calcutta, India: Pergamon Press and Statistical Publishing Society, pp.320–332.

Sun, J. (2006), *The Statistical Analysis of Interval-Censored Failure Time Data*, New York: Springer.

Vardi, Y. (1982), "Nonparametric Estimation in the Presence of Length Bias," *The Annals of Statistics*, 10, 616–620.

——— (1985), "Empirical Distributions in Selection Bias Models," *The Annals of Statistics*, 13, 178–203.

Wang, S., Liang, Z., Wang, B., and Zhang, C. (2006), "Statistical Characterization of Single-Wall Carbon Nanotube Length Distribution," *Nanotechnology*, 17, 634–639.

Ziegler, K. J., Gu, Z., Peng, H., Flor, E. L., Hauge, R. H., and Smalley, R. E. (2005), "Controlled Oxidative Cutting of Single-Walled Carbon SENTs," *Journal of the American Chemical Society*, 127, 1541–1547.

Multi-Causes Degradation Path model: A case study on rubidium lamp Degradation

Sun Quan[a,*], Paul H. Kvam [b]

[a] Department of System Engineering, College of Information System and Management, National University of Defense Technology, Changsha, HN 410073, China

[b] H. Milton Stewart School of Industrial & System Engineering, Georgia Institute of Technology, Atlanta, GA 30318, USA

Abstract

At the core of satellite rubidium standard clocks is the rubidium lamp, which is a critical piece of equipment in a satellite navigation system. There are many challenges in understanding and improving the reliability of the rubidium lamp, including the extensive lifetime requirement and the dearth of samples available for destructive life tests. Experimenters rely on degradation experiments to assess the lifetime distribution of highly reliable products that seem unlikely to fail under the normal stress conditions, because degradation data can provide extra information about product reliability. Based on recent research on the rubidium lamp, this article presents a multi-causes degradation path model, including its application background, model description, modeling method, and parameter estimation method. Using available data from degradation tests, we construct point estimates and interval estimates for rubidium lamp lifetimes using regression techniques.

Keywords

Multi-cause degradation path model, rubidium lamp, reliability

* Corresponding author.

E-mail address: sunquan@nudt.edu.cn

1. Introduction

A state-of-the-art global navigation satellite system, like GPS, GALILEO, GLONASS, can provide a highly accurate, guaranteed global positioning services for industries, governments and individuals around the world. Atomic clocks represent critical components of the satellite navigation system. An atomic clock is a sophisticated timekeeping device that uses an atomic resonance frequency standard as its timekeeping element. They are the most accurate time and frequency standards known, and are used as primary standards for international time distribution services, to control the frequency of television broadcasts, and in global navigation satellite systems such as GPS. The Rubidium Atomic Frequency Standard (RAFS) is at present the baseline clock technology for the global navigation satellite payload. According to this baseline, every satellite will embark at least two RAFSs. The adoption of a "dual technology" for the on-board clocks is dictated by the need to insure a sufficient degree of reliability (technology diversity) and to comply with the Galileo [18] lifetime requirement (12 years) [5].

1.1. Background

The rubidium lamp is the core of the RAFS and other optically pumped devices. The conventional lamp used in the RAFS consists of a glass envelope containing a charge of the appropriate metal such as rubidium, and a buffer gas under a pressure of a few torr. The lamp is usually ignited by an RF coil that surrounds the glass envelope. The spectral mission properties of the lamp have long been recognized as extremely important in the operation of optically pumped devices, and have been thoroughly studied [1]. When the rubidium atoms change energy levels for the RF discharge power, electrons in atoms will emit the precise microwave signal as the source of standard frequencies. Figure 1 shows a typical rubidium discharge lamp. The lamp consists of a glass envelope which contains the excess rubidium metal and a buffer gas. An RF coil (not shown), surrounds the exposed portion of the glass envelop and sustains a plasma in the lamp.

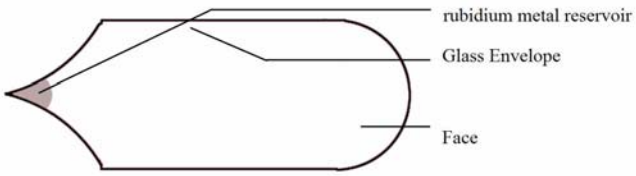


Figure 1. Typical rubidium discharge lamp

Only a limited amount of rubidium can be placed in the lamp due to concerns over the stability of its output intensity. If amount of rubidium contained in the lamp is depleted past a given threshold, the rubidium lamp will fail to operate. Several possible lamp failure mechanisms were investigated in response to the failures of the rubidium atomic clocks on board GPS satellites [2]. These mechanisms included the quenching of the excited rubidium atoms, rubidium reaction with impurities and the interaction of rubidium with the glass. Investigations reveal that the interaction of rubidium with the glass and the rubidium reaction with impurities are the two main reasons of rubidium depletion, and the interaction between the rubidium and the glass had depleted most of the rubidium [3].

1.2. The interaction of rubidium with the glass

A mathematical model [1] of the interaction of rubidium with the glass had been given by

$$Z(t) = 2 \cdot A \cdot C_0 \cdot (\frac{\pi}{D}) \cdot \sqrt{t}, t > 0, \quad (1)$$

where A is the rubidium lamp surface area, C_0 is the density of the penetrating species at the glass surface, D is the diffusion coefficient of the penetrating species for the particular glass, $Z(t)$ represents the total rubidium interaction amount to time t. The difficulty with Eq. (1) exists in obtaining precise knowledge of both C_0 and D in the lamp application. The density of rubidium at the glass surface certainly depends on the temperature of the rubidium reservoir that controls the rubidium vapor pressure in the lamp [4]. However, based on a mathematical model of the discharge lamp, the density of rubidium at the lamp wall is also predicted to be a strong function of the RF drive power applied to the lamp.

Unfortunately, at present it is not possible to model the exact functional dependence of rubidium density on lamp drive power. Additionally, the diffusion coefficients for rubidium in glasses of interest are not well known. The analyses that we have performed on various lamps have not provided sufficiently detailed information with regard to the diffusion coefficients. The best way, to date, to determine the rate of rubidium diffusion into the glass envelope of a lamp is by performing rubidium depletion measurements under certain lamp conditions of interest [3].

For convenience, we can write

$$Z(t) = b_0 + a_0 \cdot \sqrt{t}, t > 0, \quad (2)$$

Where b_0 represents the initial interaction amount, a_0 is equal to $2AC_0\pi/D$ in Eq. (1).

1.3. Rubidium reaction with Impurities

Another postulated mechanism for the failure of the rubidium lamp was the loss of rubidium by reaction with species outgassed from the envelope to form non-volatile rubidium oxide (Rb_2O). The most likely reactions are:

$$2Rb + H_2O \rightarrow Rb_2O + H_2,$$
$$4Rb + O_2 \rightarrow 2Rb_2O.$$

Through these reactions, 30 or more micrograms rubidium will be depleted. Although this amount of depletion is not significant in terms of the total charge of rubidium (300 micrograms or more), it should be taken into account to properly analysis the rubidium lamp's lifetime and its reliability.

A mathematical model [1] of the rubidium reaction with impurities can be given by

$$W(t) = P \cdot (1 - e^{-Q \cdot t}), t > 0, \quad (3)$$

where P represents the total amount of rubidium lost through reaction, Q is the reaction rate constant [2], and $W(t)$ represents the total rubidium reaction amount to time t. From Eq. (3), we know that $W(t)$ will tend to be infinitely close to a maximum amount (P) and the reaction with impurities can be considered as stopping when t is long enough, so another mathematical model of the rubidium reaction with impurities can be given by

$$W(t) = b_1 + a_1 t, \qquad t < T_1, \tag{4}$$

where b_1 represents the initial reaction amount and a_1 is the average reaction rate before the reaction stops by time T_1.

2. Rubidium degradation

The lifetime of a rubidium lamp is closely related to the amount of rubidium atoms in lamp [6, 7]. A rubidium atom's interaction with the lamp envelope glass (diffusion into the lamp's glass wall) and its reaction with impurities within the lamp will lead to the depletion of rubidium. When the amount of rubidium atom is depleted, the lamp's optical pumping will fail and the lamp fails.

Because of the challenges in obtaining lifetime data for rubidium lamps, the ability to analyze degradation measurements is paramount in accurately estimating lifetime distributions. The rubidium lamp is used only in some highly specialized equipment, such as satellites. Consequently, the demand for these lamps is low though it serves a critical purpose. In this case, the rubidium lamp is made in special laboratories, and due to manufacturing and economic constraints, there are few rubidium lamps that can used in destructive life tests.

Recently, degradation data have been shown to be a superior alternative to lifetime data in most statistical analysis because they are potentially more informative. There are two major aspects of modeling for degradation data [9]. One approach is to assume that the degradation is a random process in time. Doksum [10] used a Wiener process model to analyze degradation data. Tang and Chang [11] modeled non-destructive accelerated degradation data from power supply units as a collection of stochastic processes. Whitmore and Shenkelberg [12] considered that the degradation process in the model is taken to be a Wiener diffusion process with a time scale transformation. An alternative approach is to consider more general statistical models. Degradation in these models is modeled by a function of time and some possibly multidimensional random variables. These models are called "general degradation path models".

Lu and Meeker [13] developed statistical methods using degradation measures to estimate a time-to-failure distribution for a broad class of degradation models. They considered a nonlinear mixed-effects model (NMLE) and used a two-stage method to obtain point estimates and confidence intervals of percentiles of the failure-time distribution. Lu et al. [14] proposed a model with random regression coefficients and standard-deviation function for analyzing linear degradation data from semiconductors. Su et al. [15] considered a random coefficient degradation model with random sample size and used maximum likelihood (ML) for parameter estimation. Hamada [16] used a Bayesian approach for analyzing a laser degradation data. Bae and Kvam [17] proposed a change-point model for modeling incomplete burn-in during the production of display devices.

This aggregate of previous research has a focus on estimating the parameters in the degradation model along with the percentiles of the failure-time distribution, and these models all

assume that degradation is due to either a single cause, or when there are multiple causes, they have an additive effect that can be modeled into a single additive model. In some cases, such assumptions are not suitable.

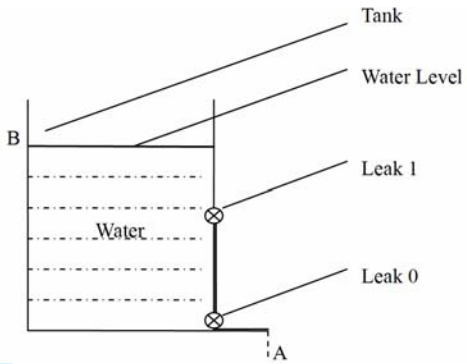


Figure 2. Water tank with Leaks at bottom (Leak 0) and middle (Leak 1).

To illustrate a case in which the regular model assumptions do not hold, consider the water tank pictured in Figure 2. When water is empty, the tank will fail to operate. From Figure2, we can see that there are two causes that lead to tank failure, labeled Leak 0 (bottom) and Leak 1 (middle). These two different causes can be analyzed using a simple degradation model that assumes a single cause exists, based on the water loss data observed from point A or tank water level falling data which can be observed from point B. However, the causes for the two Leaks are different. Leak 1 will stop when the water level is lower than a particular level, but the Leak 0 will not stop until the water completely gone.

Let $d_1(t)$ and $d_0(t)$, be the amount of Leaking water Leaking at time t corresponding to Leak 1 and Leak 0. $Y(t)$, defined as the total water Leaking amount at time t, can be described as

$$Y(t) = d_1(t) + d_0(t), \qquad t > 0. \tag{5}$$

Suppose $d_1(t)$ and $d_0(t)$ are linear functions of t, so that the rate of water Leaking from Leak 1 and Leak 0 is constant. Then,

$$d_1(t) = a_1 t + b_1, 0 \le t \le T_1, \tag{6}$$

$$d_0(t) = a_0 t + b_0, 0 \le t. \tag{7}$$

T_1 is the time at which Leak 1 ceases, when the water level is so low that there will be no water leaking from Leak 1. Then from Eq. (5) to (7), we can get a crooked line as shown in Figure 3.

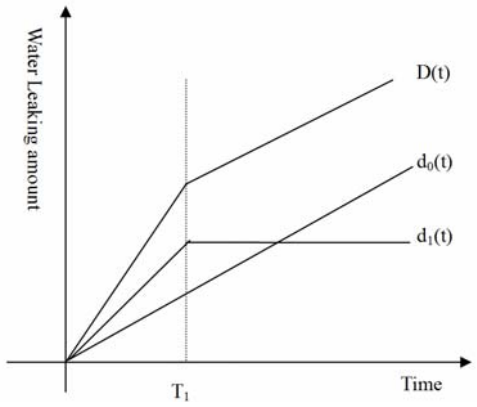


Figure 3. Amount of tank water Leakage as a function of time

In fact, the degradation of the rubidium lamp can be compared to the loss of water in tank. In the lamp, there are also two disparate causes that lead to the depletion of rubidium. One cause is the reaction with impurities, but such a reaction will stop when the impurities are used up, which occurs long before the lamp fails. The other cause is the interaction with glass (diffusion), and this kind of diffusion will continue until all rubidium completely depleted.

Bae and Kvam [17] proposed a model with change-point to analyze a similar problem in light displays. In the change-point model, several causes actually are integrated into one cause (path) and the degradation process is modeled as two independent phases between the change-point. Especially in accelerated degradation tests, the degradation rate can be modified by changes in stress levels (such as temperature) and is probably different from cause to cause, so we would study these causes separately.

In this article, we call such tank degradation problem or rubidium degradation problem as a multi-causes degradation problem, and we will develop a generalized multi-causes degradation path model to solve the multi-causes degradation problem.

3. Multi-causes degradation Path model

For convenience, we will consider two kinds of causes in our generalized multi-causes degradation path model. One cause is limited degradation. When the limit is reached, the degradation is halted, like the way Leak 1 stops once the water in the tank goes below a fixed threshold. The other cause is unlimited degradation, meaning that the degradation process will persist until failure occurs. Leak 0 in the water tank is an example of unlimited degradation. Because the continuity and similarity of independent sources of unlimited degradation, we can integrate all unlimited degradation processes into a single process. But for limited degradation processes, we can not simply integrate them into one process if the thresholds are not identical.

Assume there are $M+1$ causes of degradation, which we will call Cause 0, Cause 1, Cause 2, …, Cause M. Cause 0 is unlimited degradation with degradation process denoted by $d_0(t)$, $t > 0$. Cause 1, Cause 2, …, Cause M represent limited degradation and we denote the limit threshold for the ith cause with Df_i, $i = 1,\ldots,M$, so that

$$d_i(t) = \begin{cases} f_i(t), 0 \le t \le T_i \\ Df_i, t > T_i \end{cases}, \tag{8}$$

where T_i $(i=1,\ldots,M)$ is the time which limitation achieved. If $D(t)$ represents the item's degradation process, then

$$D(t)=\sum_{i=0}^{M} d_i(t)\,. \tag{9}$$

According to the definition in [19], we can express performance reliability as

$$\begin{aligned}
R(t)&=\Pr\{D(t)<D_f\}=\Pr\{\sum_{i=0}^{M} d_i(t)<D_f\}=\Pr\{d_0(t)<D_f-\sum_{i=1}^{M} f_i(t)\}\cdot\prod_{i=1}^{M}(1-F_i(t))\\
&+\Pr\{d_0(t)<D_f-\sum_{i=1}^{M-1} f_i(t)-D_{f_M}\}\cdot\prod_{i=1}^{M-1}(1-F_i(t))\cdot F_M(t)\\
&+\Pr\{d_0(t)<D_f-\sum_{i=1}^{M-2} f_i(t)-\sum_{i=1}^{2} D_{f_i}\}\cdot\prod_{i=1}^{M-2}(1-F_i(t))\cdot\prod_{i=1}^{2} F_{M-i+1}(t)\\
&+\ldots+\Pr\{d_0(t)<D_f-\sum_{i=1}^{M-k} f_i(t)-\sum_{i=1}^{k} D_{f_i}\}\cdot\prod_{i=1}^{M-k}(1-F_i(t))\cdot\prod_{i=1}^{k} F_{N-i+1}(t)\\
&+\ldots+\Pr\{d_0(t)<D_f-\sum_{i=1}^{M} D_{f_i}\}\cdot\prod_{i=1}^{M} F_{M-i+1}(t)
\end{aligned} \tag{10}$$

where $F_i(t)$ is the distribution function of T_i, $i = 1,\ldots, M$.

4. Rubidium Lamp Life Test

During the past several years, in order to figure out the necessary requirements that will ensure reliable operation of the rubidium lamp, there have been a number of experiments undertaken by the China State Key Laboratory of Magnetic Resonance & Atomic & Molecular Physics. The lab collected various measurements to quantify the quality of the lamps, including degradation data which we feature in this article.

In one study, there were 6 lamps of the same type of glass put into degradation test. The first three lamps were tested for 28800 hours (3 years and 4 months). Figure 4 shows the measured amount of rubidium in lamps #1 - #3. Note that the time measurement intervals are not identical. The next three lamps were tested with the same operating conditions as first three, and testing began at the same date. Only after 1 year and 8 months of regular use, degradation measurements were made for the second group of lamps, which lasted for another 14400 hours (1 year and 8 months). Figure 5 shows the rubidium amount in lamps #4 - #6, and again the time measurement intervals are not equal. Because these six lamps had worked under same usage conditions, and all started work at the same time, we can combine the degradation test data as shown in Figure 6.

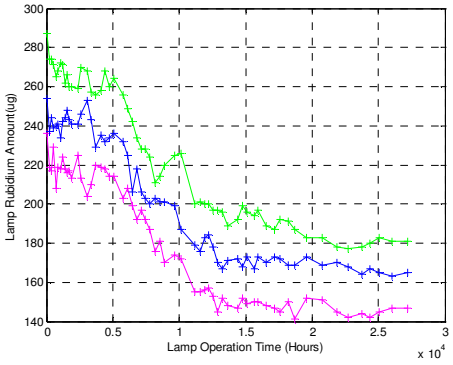


Figure 4. Rubidium Amount of Lamps #1-#3 (RF power = 2W)

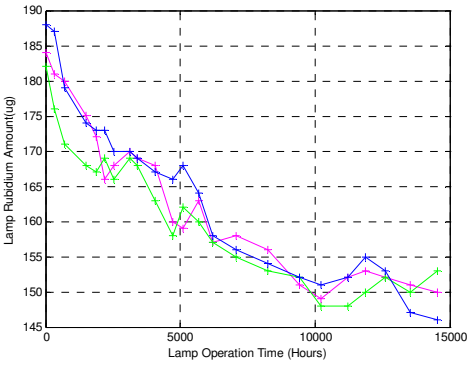


Figure 5. Rubidium Amount of Lamps #4-#6 (RF power = 2W)

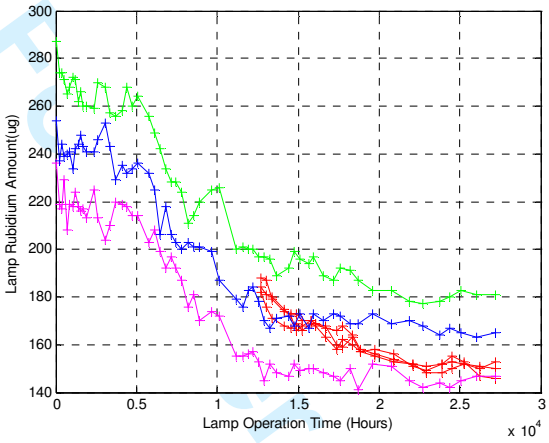


Figure 6. Rubidium Amount of Lamps #1-#6 (RF power = 2W)

5. Multi-causes degradation path model of rubidium lamp

We know that the amount of rubidium in the rubidium lamp is affected by the initial amount of rubidium that exists in the lamp, as well as the reaction rate of the impurities and the glass interaction rate. First, in the satellite rubidium atomic clock, considering the rubidium lamp physical output's stability, there is a maximal initial fill amount that varies slightly for each lamp. Although the reaction between the rubidium and the impurities continues until the lamp stops working, this depletion occurs early in the lamp lifetime, and after a particular time, the rubidium amount depleted by the reaction of rubidium and impurities becomes negligible. In contrast, degradation due to the glass interaction will continue from beginning to the end of lamp life.

5.1. Parameter evaluating method

For the rubidium lamp, there are two causes ($M+1=2$) leading to the degradation. One major cause is due to interaction with the glass, and its degradation path can be denoted as following from Eq. (2):

$$d_0(t) = Z(t) = b_0 + a_0 \cdot \sqrt{t}, t > 0 \qquad (11)$$

The other degradation cause is due to the reaction with impurities, and its degradation path can be

denoted as following from Eq. (4):

$$d_1(t)=\begin{cases}W(t),0\le t\le T_1\\ Df_1,t>T_1\end{cases}=\begin{cases}b_1+a_1\cdot t,0\le t\le T_1\\ Df_1,t>T_1\end{cases} \tag{12}$$

From Eq. (9), we obtain (M=1):

$$D(t)=\sum_{i=0}^{M}d_i(t)=d_0(t)+d_1(t)=\begin{cases}b_0+a_0\cdot\sqrt{t}+b_1+a_1\cdot t,0\le t\le T_1\\ b_0+a_0\cdot\sqrt{t}+Df_1,t>T_1\end{cases}$$
$$=\begin{cases}(b_1+b_0)+a_0\cdot\sqrt{t}+a_1\cdot t,0\le t\le T_1\\ (b_0+Df_1)+a_0\cdot\sqrt{t},t>T_1\end{cases} \tag{13}$$

For convenience, let

$$Y(t)=D_f-D(t),c=D_f-b_1-b_0,d=D_f-b_0-Df_1, \tag{14}$$

then Eq. (13) can be written as:

$$Y(t)=\begin{cases}c-a_1t-a_0\sqrt{t},0\le t\le T_1\\ d-a_0\sqrt{t},t>T_1\end{cases}, \tag{15}$$

where $Y(t)$ is a continuous function of time, c represents the rubidium initial fill amount parameter, a_1 represents rubidium and impurities reaction parameter, a_0 represents the rubidium and glass interaction parameter, T_1 represents intersection point which limitation cause stops operating, and c, a_1, a_0, d, T_1 are all positive real numbers.

If $[t_{i-1}, t_i]$ represents the ith time interval, $i = 1,\ldots,N$ (with $t_0 = 0$, N is the test time intervals), and $Y(t_i)$ is the measured amount of rubidium in the lamp at time t_i, we can get following equation from Eq. (15):

$$Y(t_i)=\begin{cases}c-a_1t_i-a_0\sqrt{t_i}+\varepsilon_1,i=1,2,\ldots,\tau\\ d-a_0\sqrt{t_i}+\varepsilon_2,t>T_1,i=\tau+1,\ldots N\end{cases}, \tag{16}$$

where (ε_1, ε_2) are the random measurement errors, assumed to be independent and identically distributed (iid) normal random variables with mean zero and variance σ^2. The parameter τ is the index corresponding to the intersection point T_1, with $\tau = 1, \ldots, N$, and $t_\tau \le T_1 \le t_{\tau+1}$.

Additionally, under the continuity assumption for the degradation path, we can get

$$T_1 = (c\text{-}d)/a_1, \tag{17}$$

where d ∈ [c - $a_1 t_{\tau+1}$, c + $a_1 t_{\tau+1}$], c ≥ 0, $a_1 \geq 0$, $a_0 \geq 0$, d ≥ 0. Using the model described in Eq. (17), it is possible to derive maximum likelihood estimators (MLEs) for c, a_1, a_0, d. The log-likelihood is given by:

$$\mathrm{L}(c,a_1,a_0,d,\sigma^2,\tau)=\ln\frac{1}{(\sqrt{2\pi})^N\sigma^N}-\sum_{i=1}^{\tau}\frac{(y(t_i)-(c-a_1t_i-a_0\sqrt{t_i}))^2}{\sigma^2}-\sum_{i=\tau+1}^{N}\frac{(y(t_i)-(d-a_0\sqrt{t_i}))^2}{\sigma^2} \quad (18)$$

The procedure to obtain MLEs for the parameters in Eq. (18) along with the corresponding estimate of τ can be summarized in a simple constrained optimization:

Find $\hat{c},\hat{a}_1,\hat{a}_0,\hat{d},\hat{\sigma}^2$, and $\hat{\tau}$ such that:

$$L(\hat{c},\hat{a}_1,\hat{a}_0,\hat{d},\hat{\sigma}^2,\hat{\tau})=\sup \mathrm{L}(c,a_1,a_0,d,\sigma^2,\tau), \quad (19)$$

subject to

$$\hat{T}_1=\frac{\hat{c}-\hat{d}}{\hat{a}_1}, \text{ and } \hat{d}\in[\hat{c}-\hat{a}_1 t_{\tau+1},\hat{c}-\hat{a}_1 t_\tau], \hat{a}_1\geq 0, \hat{a}_0\geq 0, \hat{d}\geq 0, \hat{c}\geq 0.$$

Under the constraints above, the MLEs $\hat{c},\hat{a}_1,\hat{a}_0,\hat{d},\hat{\sigma}^2$ can be found in closed form:

$$\hat{c}=\{\frac{T\cdot Y_\tau\cdot(N-\tau)+T_{N-\tau}^{\frac{1}{2}}\cdot T_{N-\tau}^{\frac{1}{2}}\cdot Y_\tau+Y_{N-\tau}\cdot T_{N-\tau}^{\frac{1}{2}}\cdot T_\tau^{\frac{1}{2}}-T_\tau^{\frac{1}{2}}\cdot YT^{\frac{1}{2}}\cdot(N-\tau)}{T_\tau^{\frac{3}{2}}\cdot T_\tau^{\frac{1}{2}}\cdot(N-\tau)-T\cdot T_\tau\cdot(N-\tau)-T_\tau\cdot T_{N-\tau}^{\frac{1}{2}}\cdot T_{N-\tau}^{\frac{1}{2}}}$$
$$-\frac{2\cdot T\cdot YT_\tau\cdot(N-\tau)+2\cdot T_{N-\tau}^{\frac{1}{2}}\cdot T_{N-\tau}^{\frac{1}{2}}\cdot YT_\tau+T_{N-\tau}^{\frac{1}{2}}\cdot T_\tau^{\frac{3}{2}}\cdot Y_{N-\tau}-T_\tau^{\frac{3}{2}}\cdot YT^{\frac{1}{2}}\cdot(N-\tau)}{T_\tau^{\frac{3}{2}}\cdot T_\tau^{\frac{3}{2}}\cdot(N-\tau)-T_\tau^2\cdot T\cdot(N-\tau)-T_\tau^2\cdot T_{N-\tau}^{\frac{1}{2}}\cdot T_{N-\tau}^{\frac{1}{2}}}\},$$
$$/\{\frac{\tau\cdot T\cdot(N-\tau)+\tau\cdot T_{N-\tau}^{\frac{1}{2}}\cdot T_{N-\tau}^{\frac{1}{2}}-T_\tau^{\frac{1}{2}}\cdot T_\tau^{\frac{1}{2}}\cdot(N-\tau)}{T_\tau^{\frac{1}{2}}\cdot T_\tau^{\frac{3}{2}}\cdot(N-\tau)-T_\tau\cdot T\cdot(N-\tau)-T_\tau\cdot T_{N-\tau}^{\frac{1}{2}}\cdot T_{N-\tau}^{\frac{1}{2}}}$$
$$-\frac{T_\tau\cdot T\cdot(N-\tau)+T_\tau\cdot T_{N-\tau}^{\frac{1}{2}}\cdot T_{N-\tau}^{\frac{1}{2}}-T_\tau^{\frac{3}{2}}\cdot T_\tau^{\frac{1}{2}}\cdot(N-\tau)}{T_\tau^{\frac{3}{2}}\cdot T_\tau^{\frac{3}{2}}\cdot(N-\tau)-T_\tau^2\cdot T\cdot(N-\tau)-T_\tau^2\cdot T_{N-\tau}^{\frac{1}{2}}\cdot T_{N-\tau}^{\frac{1}{2}}}\},$$

$$\hat{a}_1 = \frac{2YT_\tau \cdot T.(N-\tau) + 2YT_\tau \cdot T_{N-\tau}^{\frac{1}{2}} \cdot T_{N-\tau}^{\frac{1}{2}} + T_\tau^{\frac{3}{2}} \cdot Y_{N-\tau} \cdot T_{N-\tau}^{\frac{1}{2}} - T_\tau^{\frac{3}{2}} \cdot YT^{\frac{1}{2}} \cdot (N-\tau)}{T_\tau^{\frac{3}{2}} \cdot T_\tau^{\frac{3}{2}} \cdot (N-\tau) - T_\tau^2 \cdot T \cdot (N-\tau) - T_\tau^2 \cdot T_{N-\tau}^{\frac{1}{2}} \cdot T_{N-\tau}^{\frac{1}{2}}},$$

$$-\frac{T_\tau \cdot T.(N-\tau) + T_\tau \cdot T_{N-\tau}^{\frac{1}{2}} \cdot T_{N-\tau}^{\frac{1}{2}} - T_\tau^{\frac{3}{2}} \cdot T_\tau^{\frac{1}{2}} \cdot (N-\tau)}{T_\tau^{\frac{3}{2}} \cdot T_\tau^{\frac{3}{2}} \cdot (N-\tau) - T_\tau^2 \cdot T \cdot (N-\tau) - T_\tau^2 \cdot T_{N-\tau}^{\frac{1}{2}} \cdot T_{N-\tau}^{\frac{1}{2}}} \cdot H$$

$$\hat{a}_0 = \frac{Y_{N-\tau} \cdot T_{N-\tau}^{\frac{1}{2}} - YT^{\frac{1}{2}} \cdot (N-\tau) + T_\tau^{\frac{1}{2}} \cdot (N-\tau) \cdot H - T_\tau^{\frac{3}{2}} \cdot (N-\tau) \cdot \hat{a}_1}{T \cdot (N-\tau) + T_{N-\tau}^{\frac{1}{2}} \cdot T_{N-\tau}^{\frac{1}{2}}},$$

$$\hat{d} = \frac{Y_{N-\tau} - T_{N-\tau}^{\frac{1}{2}} \cdot \hat{a}_0}{(N-\tau)},$$

$$\hat{\sigma}^2 = \frac{2}{N}\left(\sum_{i=1}^{\tau}(y(t_i) - (c - \hat{a}_1 \cdot t_i - \hat{a}_0 \cdot \sqrt{t_i}))^2 + \sum_{i=\tau+1}^{N}(y(t_i) - (\hat{d} - \hat{a}_0 \cdot \sqrt{t_i}))^2\right), \quad (20)$$

where

$$T = \sum_{i=1}^{N} t_i, T_\tau = \sum_{i=1}^{\tau} t_i, T_{N-\tau} = \sum_{i=\tau+1}^{N} t_i,$$

$$T^{\frac{1}{2}} = \sum_{i=1}^{N} t_i^{\frac{1}{2}}, T_\tau^{\frac{1}{2}} = \sum_{i=1}^{\tau} t_i^{\frac{1}{2}}, T_{N-\tau}^{\frac{1}{2}} = \sum_{i=\tau+1}^{N} t_i^{\frac{1}{2}},$$

$$T^{\frac{3}{2}} = \sum_{i=1}^{N} t_i^{\frac{3}{2}}, T_\tau^{\frac{3}{2}} = \sum_{i=1}^{\tau} t_i^{\frac{3}{2}}, T_{N-\tau}^{\frac{3}{2}} = \sum_{i=\tau+1}^{N} t_i^{\frac{3}{2}},$$

$$T^2 = \sum_{i=1}^{N} t_i^2, T_\tau^2 = \sum_{i=1}^{\tau} t_i^2, T_{N-\tau}^2 = \sum_{i=\tau+1}^{N} t_i^2,$$

$$Y = \sum_{i=1}^{N} y(t_i)^{\frac{1}{2}}, Y_\tau = \sum_{i=1}^{\tau} y(t_i), Y_{N-\tau} = \sum_{i=\tau+1}^{N} y(t_i),$$

$$YT^{\frac{1}{2}} = \sum_{i=1}^{N} y(t_i) \cdot t_i^{\frac{1}{2}}, YT_\tau = \sum_{i=1}^{\tau} y(t_i) \cdot t_i$$

$$H = \frac{T \cdot Y_\tau \cdot (N-\tau) + T_{N-\tau}^{\frac{1}{2}} \cdot T_{N-\tau}^{\frac{1}{2}} \cdot Y_\tau - (N-\tau) \cdot T_\tau^{\frac{1}{2}} \cdot YT^{\frac{1}{2}} + Y_{N-\tau} \cdot T_{N-\tau}^{\frac{1}{2}} \cdot T_\tau^{\frac{1}{2}}}{T_\tau^{\frac{3}{2}} \cdot T_\tau^{\frac{1}{2}} \cdot (N-\tau) - (N-\tau) \cdot T_\tau \cdot T - T_{N-\tau}^{\frac{1}{2}} \cdot T_{N-\tau}^{\frac{1}{2}} \cdot T_\tau}$$

$$-\frac{2T \cdot YT_\tau \cdot (N-\tau) + 2T_{N-\tau}^{\frac{1}{2}} \cdot T_{N-\tau}^{\frac{1}{2}} \cdot YT_\tau - (N-\tau) \cdot T_\tau^{\frac{3}{2}} \cdot YT^{\frac{1}{2}} + Y_{N-\tau} \cdot T_{N-\tau}^{\frac{1}{2}} \cdot T_\tau^{\frac{3}{2}}}{T_\tau^{\frac{3}{2}} \cdot T_\tau^{\frac{3}{2}} \cdot (N-\tau) - (N-\tau) \cdot T_\tau^2 \cdot T - T_{N-\tau}^{\frac{1}{2}} \cdot T_{N-\tau}^{\frac{1}{2}} \cdot T_\tau^2}.$$

Using $\hat{c}, \hat{a}_1, \hat{a}_0, \hat{d}, \hat{\sigma}^2$, and $\hat{\tau}$ in Eq. (20), from Eq. (14), we can also get the estimators for b_0, b_1 as follows:

$$\hat{b}_0 = D_f - D_{f_1} - \hat{d},$$
$$\hat{b}_1 = \hat{d} + Df_1 - \hat{c}. \quad (21)$$

5.2. Parameter evaluation

With a sample of *L* lamps, we can obtain MLEs, along with corresponding estimates their standard errors, for each lamp, treating the parameters as fixed affects unique to each lamp. From Eq. (20) and Eq. (21), we would have a set of 5L parameters, assuming that the *j*th lamp's degradation parameters are $a_{0,j}, b_{0,j}, a_{1,j}, b_{1,j}, T_{1,j}$, *j*=1,2,…,*L*.

With the data from the lamp study, using the degradation data shown in Figure 4 - 6, we obtain the multi-causes degradation path model parameter estimates and the standard errors for each individual lamp (listed in Table 1), where $L = 6$ and $Df_1 = 60\,\mu g$. For details about how the standard errors of nonlinear parameters can be efficiently computed, see reference [20].

Table 1 The MLEs and their standard errors for parameters in the rubidium lamps' degradation model

	$b_{1,j}$	$a_{1,j}$	$b_{0,j}$	$a_{0,j}$	$T_{1,j}$	σ_j^2
j = 1	0.44	0.00456	69.27	0.09779	15270.1	88.8
SE	3.867	0.001334	2.4384	0.01544		
j = 2	7.09	0.00326	9.58	0.25628	16220.2	99.14
SE	2.5726	0.001401	6.279	0.012928		
j = 3	5.94	0.00372	37.3	0.23513	14547.5	84.17
SE	3.8502	0.001063	1.3947	0.013793		
j = 4	2.15	0.00363	18.49	0.44789	15956.9	12.47
SE	0.17275	0.000376	2.7654	0.005813		
j = 5	22.95	0.00224	26.24	0.40319	16561.6	14.77
SE	0.14349	0.000766	2.8393	0.007648		
j = 6	2.05	0.00374	0.06	0.57030	15475.6	11.11
SE	0.53403	0.002852	0.64325	0.028464		

To see if the parameter estimates follow the normal distribution, we used two goodness of fit tests on the sets of 6 estimated coefficients: the Lilliefor's test and the Anderson-Darling test. Treated as a sample data of random coefficients b_1, ($b_{0,j}$, *j*=1,2,…,*6*), both tests indicate an adequate fit to a normal distribution at the 5% level. The other coefficients ($a_{0,j}$, *j*=1,2,…,*6*), ($a_{1,j}$, *j*=1,2,…,*6*), ($b_{1,j}$, *j*=1,2,…,*6*) , ($T_{1,j}$, *j*=1,2,…,*6*) also passed this Lilliefors test for goodness of fit to a normal distribution at the 5% level.

Unlike these other coefficients, T_1 represents the end (lifetime) of Rubidium reaction with impurities in the Multi-Cause Rubidium degradation path model. Based on the results of Lilliefors test of an exponential distribution, ($T_{1,j}$, j=1,2,…,6) is adequately fit by the exponential distribution at the 5% level. The resulting model now considers

$$a_0 \sim N(\mu_{a_0}, \sigma^2_{a_0}), b_0 \sim N(\mu_{b_0}, \sigma^2_{b_0}), a_1 \sim N(\mu_{a_1}, \sigma^2_{a_1}),$$
$$b_1 \sim N(\mu_{b_1}, \sigma^2_{b_1}), T_1 \sim E(\lambda_1),$$

and we seek to calculate the estimates of $\mu_{a_0}, \sigma^2_{a_0}, \mu_{b_0}, \sigma^2_{b_0}, \mu_{a_1}, \sigma^2_{a_1}, \mu_{b_1}, \sigma^2_{b_1}, \lambda_1$.

By substituting the parameter estimates in Table 1 into Eq. (13), the model generates the degradation path of each rubidium lamp, as shown from Figure 7 to Figure 12.

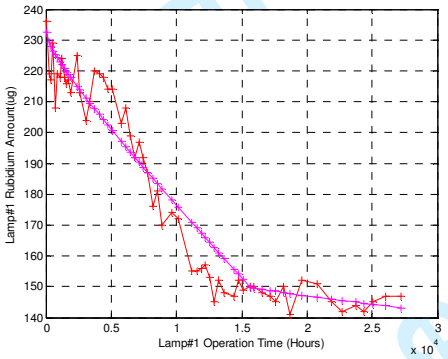


Figure 7. The degradation paths of rubidium lamps #1

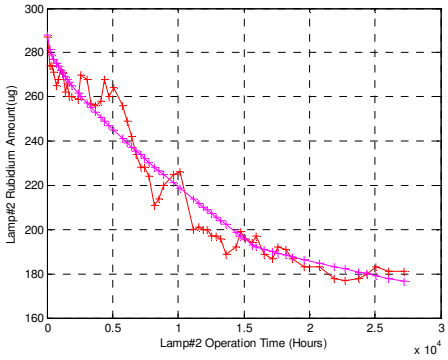


Figure 8. The degradation paths of rubidium lamps #2

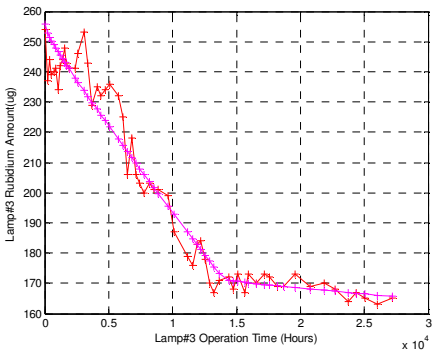


Figure 9. The degradation paths of rubidium lamps #3

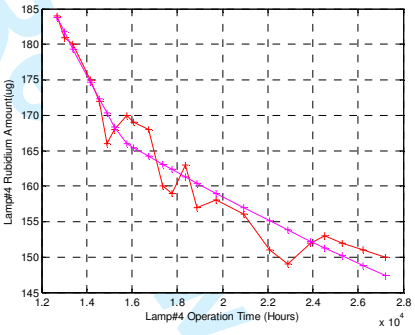


Figure 10. The degradation paths of rubidium lamps #4

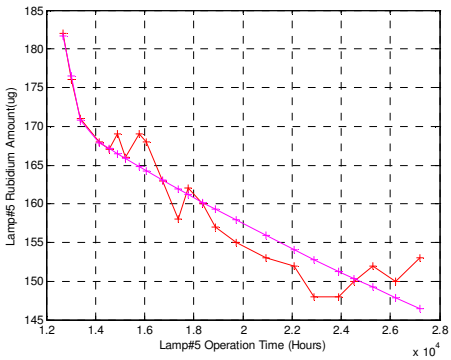


Figure 11. The degradation paths of rubidium lamps #5

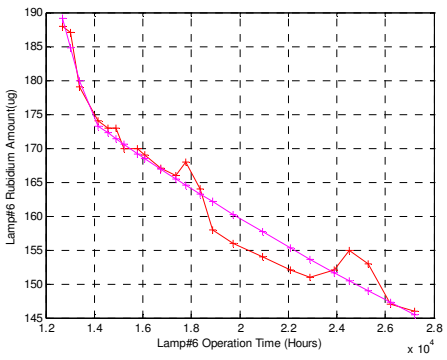


Figure12. The degradation paths of rubidium lamps #6

6. Reliability Analysis of Rubidium Lamp

In this section, we will use the degradation data to analyze the reliability of rubidium lamps. From Eq. (10), when $M = 1$, we can obtain the reliability performance of the rubidium lamp as follows:

$$\begin{aligned} R(t) &= \Pr\{D(t) < D_f\} \\ &= \Pr\{d_0(t) + d_1(t) < D_f\} \\ &= \Pr\{a_0\sqrt{t} + b_0 + a_1 \cdot t + b_1 < D_f\} \cdot (1 - F_1(t)) \\ &+ \Pr\{a_0\sqrt{t} + b_0 < D_f - D_{f_1}\} \cdot F_1(t) \end{aligned} \tag{22}$$

From Eq. (22), D_f and D_{f1} represent relative thresholds. For the rubidium lamp in this study, D_f should be set to 300 μg and D_{f1} can be set to 60 μg [1][2]. Then the rubidium lamp's performance reliability is calculated as:

$$\begin{aligned} R(t) &= \Phi\left(\frac{D_f - (\mu_{a_0}\sqrt{t} + \mu_{a_1}t + \mu_{b_0} + \mu_{b_1})}{\sqrt{\sigma_{a_0}^2 t + \sigma_{a_1} t^2 + \sigma_{b_0}^2 + \sigma_{b_1}^2}}\right) e^{-\lambda_1 t} \\ &+ \Phi\left(\frac{D_f - D_{f_1} - (\mu_{a_0}\sqrt{t} + \mu_{b_0})}{\sqrt{\sigma_{a_0}^2 t + \sigma_{b_0}^2}}\right)(1 - e^{-\lambda_1 t}) \end{aligned}, \tag{23}$$

where Φ is the CDF of the Standard Normal Distribution.

To get the performance reliability point estimation at time t, we first compute estimates of $\mu_{a_0}, \sigma_{a_0}^2, \mu_{b_0}, \sigma_{b_0}^2, \mu_{a_1}, \sigma_{a_1}^2, \mu_{b_1}, \sigma_{b_1}^2, \lambda_1$, and then substitute these estimates into Eq. (23). When D_f =300 μg, D_{f_1} =60 μg, t = 87,600 hours (10 years), using a nonparametric bootstrap resampling method, we draw 10000 bootstrap samples from the data set of Table 1. Next, we compute statistical estimates for each sample using Eq. (23), and return the results, which are illustrated in Figure 13. From the bootstrap samples, we use the 0.025 and 0.975 quantiles to form the confidence interval for the reliability estimate. For example, at 10 years (see Figure 14) we compute $R_{0.025} = 0.92579$ and $R_{0.975} = 0.99666$.

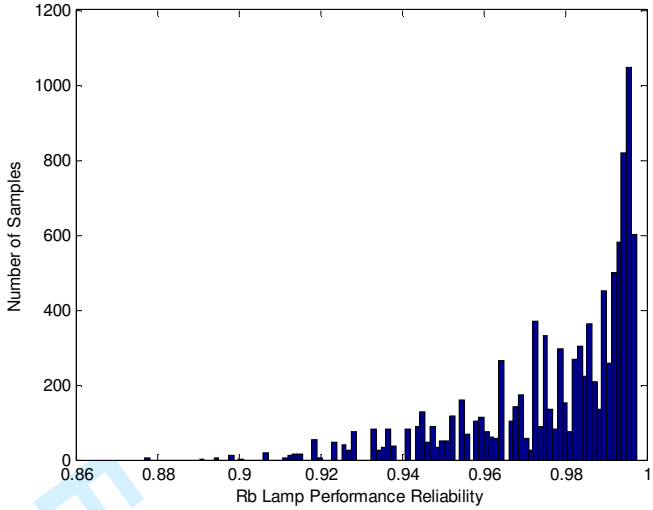


Figure 13. Rubidium Lamp Performance Reliability's Histogram by bootstrap data samples

Applying this resampling technique at various operating times, *R(t)* and its 95% bootstrap confidence intervals are plotted in Figure 14. The intervals are based on 10000 bootstrap samples, the upper curve is *R(t)* estimation for *quantile = 0.025*, the lower curve is *R(t)* estimation for *quantile = 0.975*, and the middle curve is the reliability estimate.

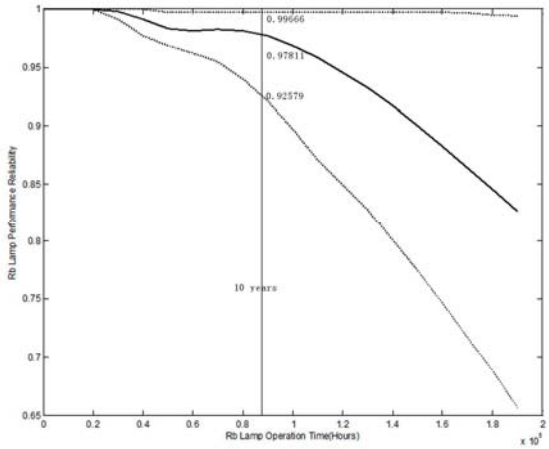


Figure 14. Estimated Performance Reliability of Rubidium Lamp Along With 95% Confidence Interval

7. Conclusion and future research

Due to the previous lack of failure data, reliability estimation for the rubidium lamp has faced great hurdles. Using the data provided by the China State Key Laboratory, this article presents multi-cause degradation path model, including its brief application background, model description, modeling method, and parameter estimation method. Using degradation data from just six life tests, we analyzed the performance reliability of the rubidium lamp and displayed the performance reliability results (point estimation and interval estimation).

As a result, this research represents the first time rubidium lamp reliability has been inferred using degradation data. We note that the results are not completely reassuring. Previous internal

lab studies based on theoretical models were comparatively optimistic about predicting lamp reliability. The reliability estimates here suggest that the rubidium lamp's performance reliability is lacking, with the lower confidence interval bound of 92% for 10 service years, which is over the tolerated failure frequency. As a matter of fact, from the rubidium lamp's multi-causes degradation path model and its performance reliability equation, we know that we need to make improvements in many aspects such as increasing the initial fill amount, decreasing the amount of impurities, and improving the glass materials, for example.

It is likely that multi-causes degradation is increasingly prevalent in degradation tests of highly complex devices, but mostly for convenience we integrated all causes into one cause and set up a degradation path that is easier to characterize. For the rubidium lamp, the high rf power will create different effects from the rubidium interaction with impurities as well as with glass, and the multi-cause degradation path model can help to recognize the difference among these effects. In future work, when more refined data become available, we will consider more specialized models of degradation that consider more causes of degradation.

Acknowledgements: This work was supported by the National Science Foundation of China under agreement 60701006 and the National Science Foundation of the United States (CMMI-0700131).

References

[1] R.A. Cook, R.P. Frueholz. An improved rubidium depletion model for discharge lamps used in rubidium Frequency standards. 42nd Annual Frequency control symposium; 1988; p.525-531.

[2] C.H. Volk, R.P. Frueholz, T.C. English, T.J. Lynch, W.J. Riley. Lifetime and reliability of rubidium discharge lamps for use in atomic Frequency standard. 38th Annual Frequency control symposium; 1984; p.387-400.

[3] IEEE Std 1193 – 1994. IEEE Guide for Measurement of Environmental Sensitivities of Standard Frequency Generators[S]. 1994.

[4] Robert P. Frueholz. The Effects of Ambient Temperature Fluctuations on the Long Term Frequency Stability of a Miniature rubidium atomic Frequency standard. IEEE INTERNATIONAL FREQUENCY CONTROL SYMPOSIUM;1996; p.1017-1022.

[5] P. Rochat, F. Droz, P. Mosset, G. Barmaverain, Q. Wang, D. Boving, et al. The onboard Galileo rubidium and passive maser, status and performance. Proceedings of 2005 Joint IEEE International Frequency and Control Symposium and Precise Time and Time Interval (PTTI) Systems and Applications Meeting; 2005; p. 26-32.

[6] Natarajan D Bhaskar. A Historical Review of Atomic Frequency Standards Used in Space Systems. Proceedings of the 1996 IEEE International Frequency Control Symposium; 1996; p. 24-32.

[7] R Beard, J Buisson, F Danzy. GPS Block IIR rubidium Frequency Standard Life Test Results. 2002 IEEE International Frequency Control Symposium and PDA Exhibition; 2002; p. 499 - 505.

[8] Chow,G. Tests of the equality between two sets of coefficients in two linear regressions. Econometrica 1960;28:561-605.

[9] Marta A. Freitas1, Maria Lu´iza G. de Toledo, Enrico A. Colosimo, Magda C. Pires. Using Degradation Data to Assess Reliability: A Case Study on Train Wheel Degradation. QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL 2009; 25:607–629.

[10] Doksum KA. Degradation rate models for failure time and survival data. CWI Quarterly 1991; 4:195–203.

[11] Tang LC, Chang DS. Reliability prediction using nondestructive accelerated degradation data: Case study on power supplies. IEEE Transactions on Reliability 1995; 44:562–566.

[12] Whitmore GA, Shenkelberg F. Modelling accelerated degradation data using Wiener diffusion with a time scale transformation. Lifetime Data Analysis 1997; 3:27–45.

[13] Lu CJ, Meeker WQ. Using degradation measures to estimate a time-to-failure distribution. Technometrics 1993;35:161–174.

[14] Lu JC, Park J, Yang Q. Statistical inference of a time-to-failure distribution from linear degradation data. Technometrics 1997; 39:391–400.

[15] Su C, Lu JC, Chen D, Hughes-Oliver JMA. Random coefficient degradation model with random sample size. Lifetime Data Analysis 1999; 5:173–183.

[16] Hamada M. Using degradation data to assess reliability. Quality Engineering 2005; 17:615–620.

[17] Bae SJ and Kvam, PH A change-point analysis approved to the modeling of incomplete burn-in during the production of display devices, IIE transactions 2006;.6:489-498.

[18] A. Jeanmaire, P. Rochat, F. Emma. RUBIDIUM ATOMIC CLOCK FOR GALILEO. 31st Annual Precise Time and Time Interval (PTTI) Meeting;1999;p.627-636.

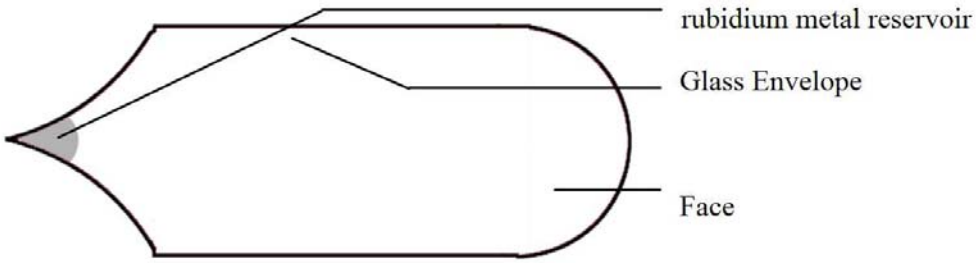
[19]William Q.Meeker, Luis A.Escobar. Statistical Methods for reliability Data. Newyork: JOHN WILEY & SONS; 2003.

[20] Douglas M. Bates, Donald G. Watts. Nonlinear Regression Analysis and Its applications. New Jersey: JOHN WILEY & SONS; 1988.

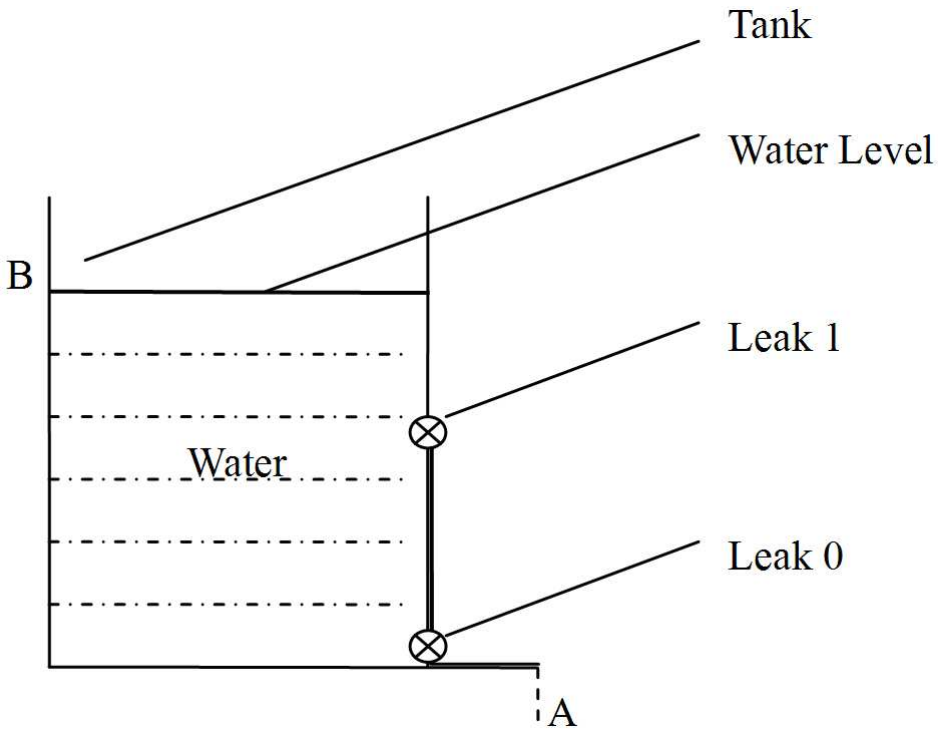
About the Authors

Sun Quan is an Associate Professor in the college of information system & management at National University of Defense Technology (NUDT), china. From March 2009 to March 2010, he worked at Georgia Institute of Technology as a visiting scholar. He received his B.S. in Mathematics from NUDT at 1994, an M.S. in Operation Research from NUDT at 1997, and his Ph.D. in System Engineering from NUDT at 2005. His research interests are in quality & reliability engineering. He is a member of IEEE.

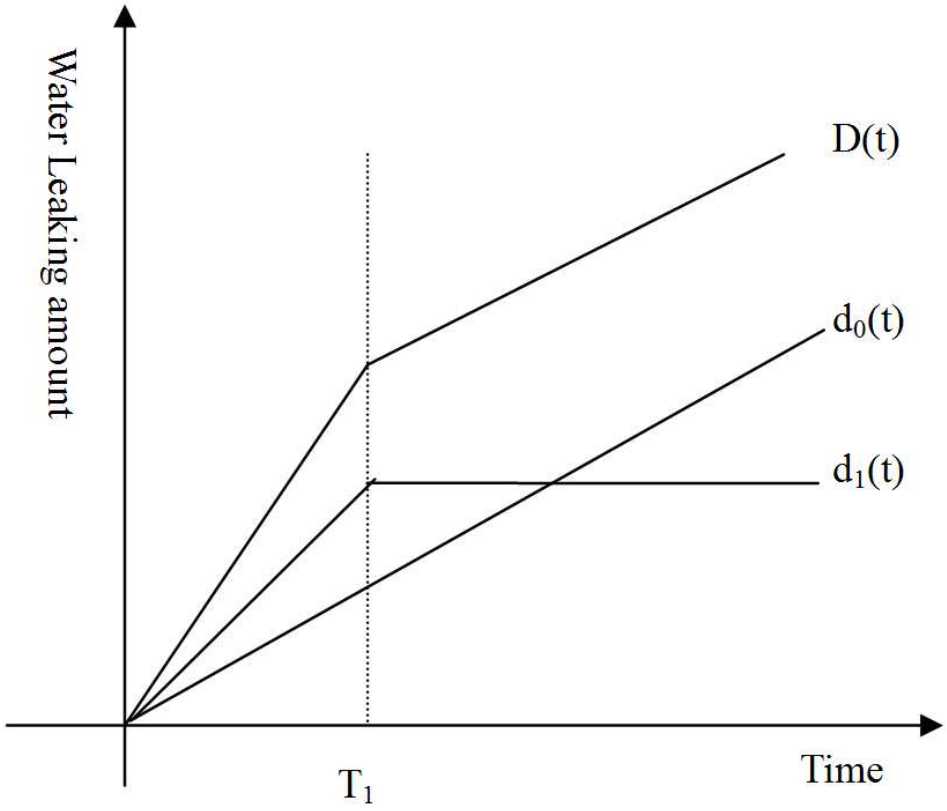
Paul Kvam is Professor in School of Industrial and Systems Engineering at Georgia Institute of Technology. He received his B.S. in Mathematics from Iowa State University in 1984, an M.S. in Statistics from the University of Florida in 1986, and his Ph.D. in Statistics from the University of California, Davis in 1991. His research interests focus on statistical reliability with applications to engineering, nonparametric estimation, and analysis of complex and dependent systems. He has served as associate editor for IEEE Transactions on Reliability (1992-2000), Technometrics (1999-2005), The American Statistician (2005-present) and Journal of the American Statistical Association (2002-present). He is a Fellow of the American Statistical Association, as well as an active member of Institute of Mathematical Statistics and Institute for Operations Research and Management Science.



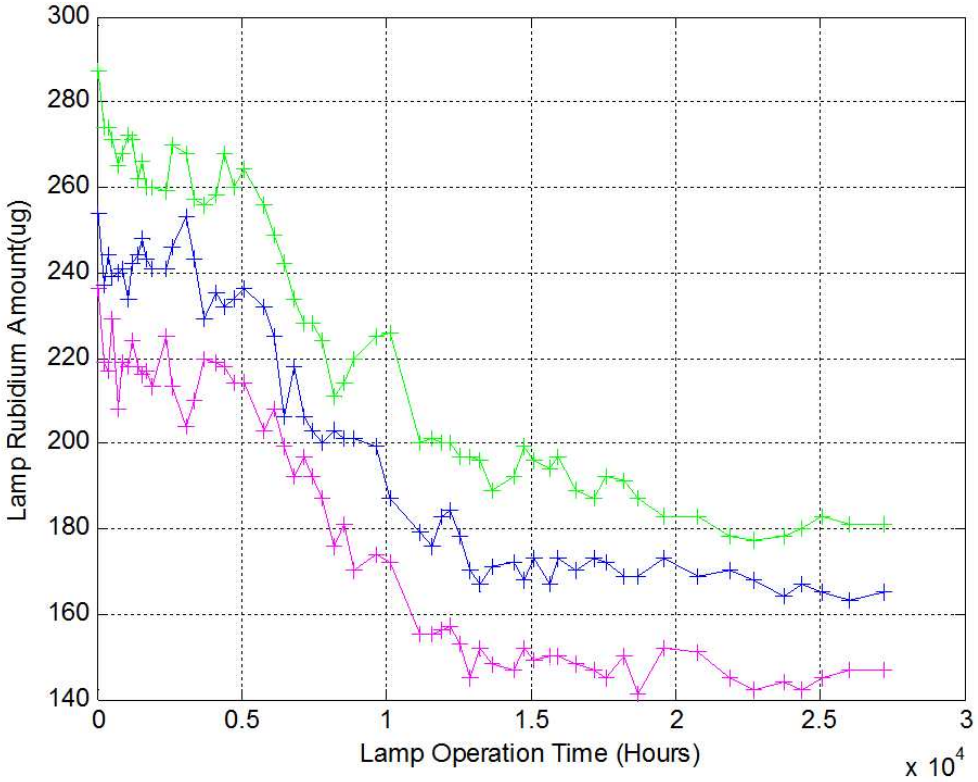
104x33mm (300 x 300 DPI)



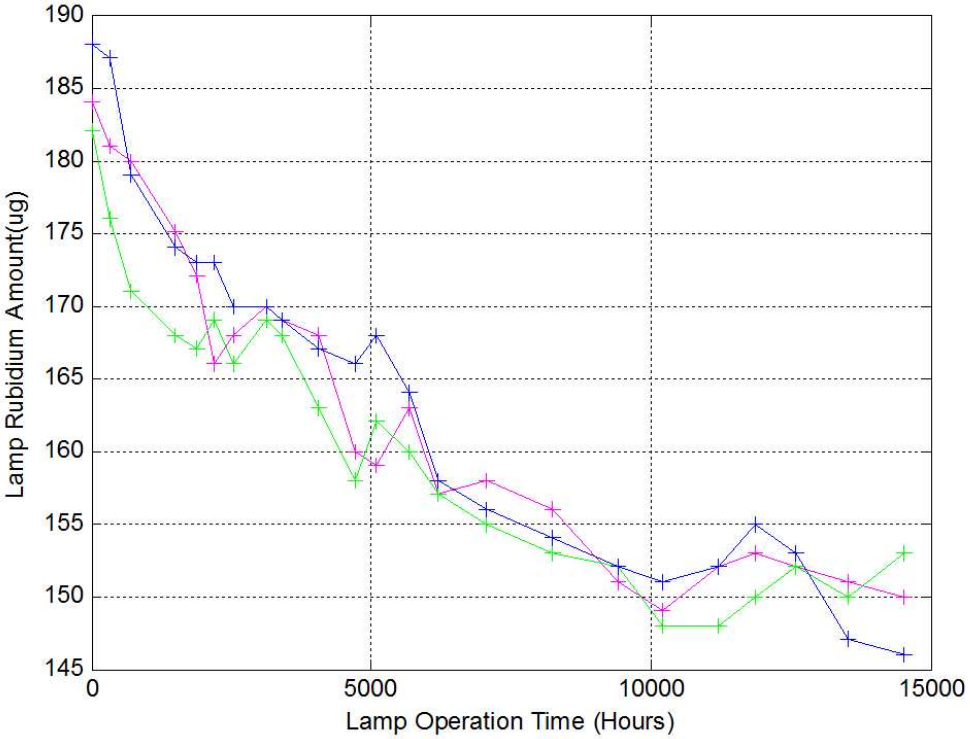
84x61mm (300 x 300 DPI)



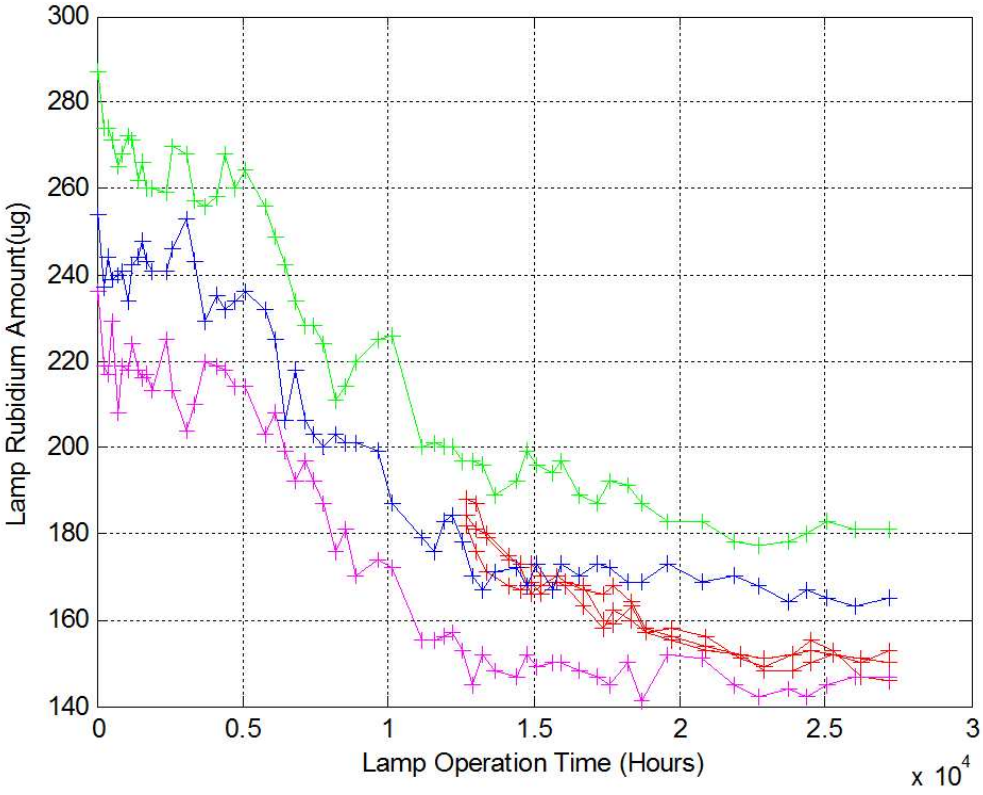
73x59mm (300 x 300 DPI)



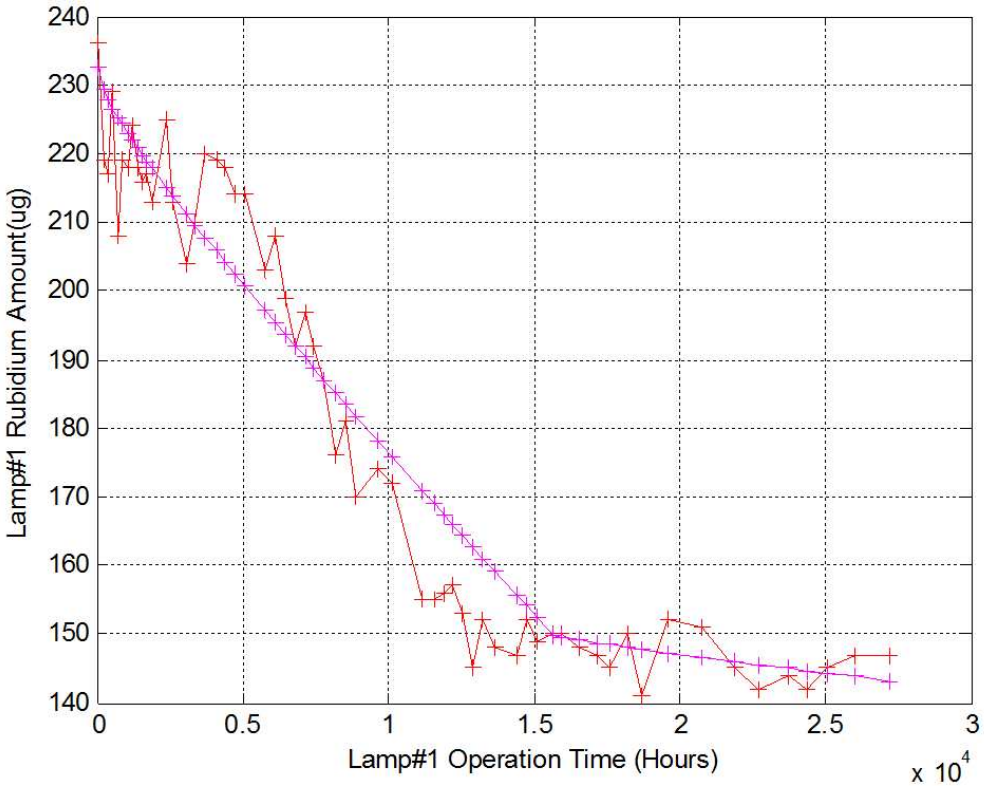
77x61mm (300 x 300 DPI)



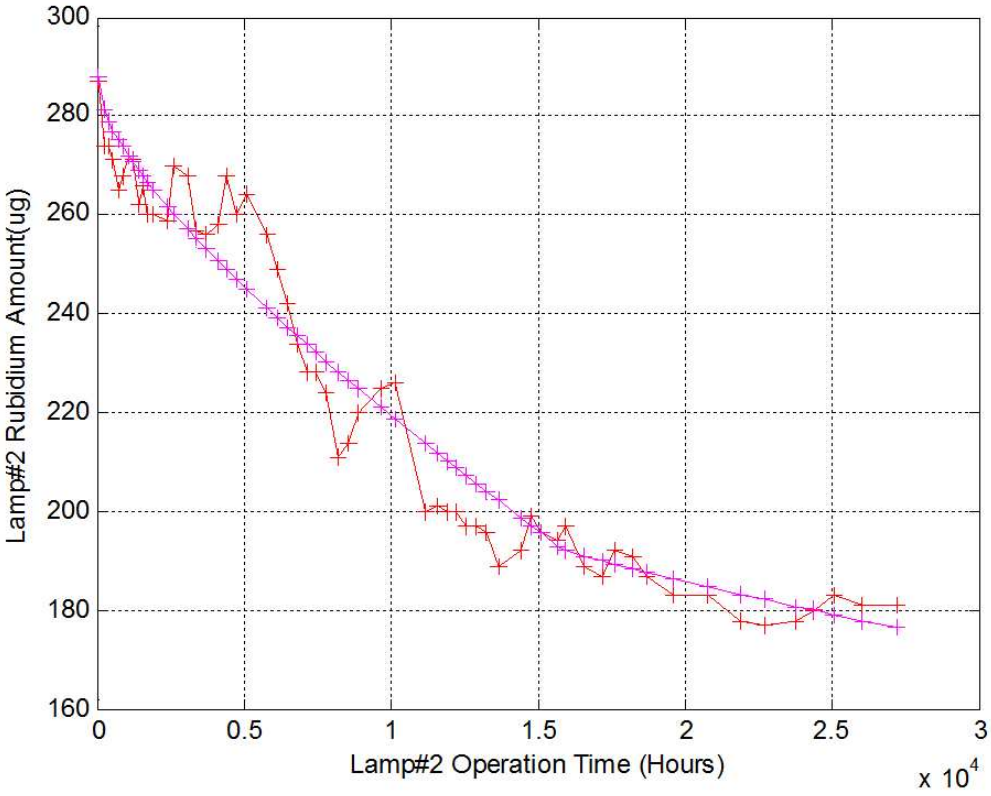
82x62mm (300 x 300 DPI)



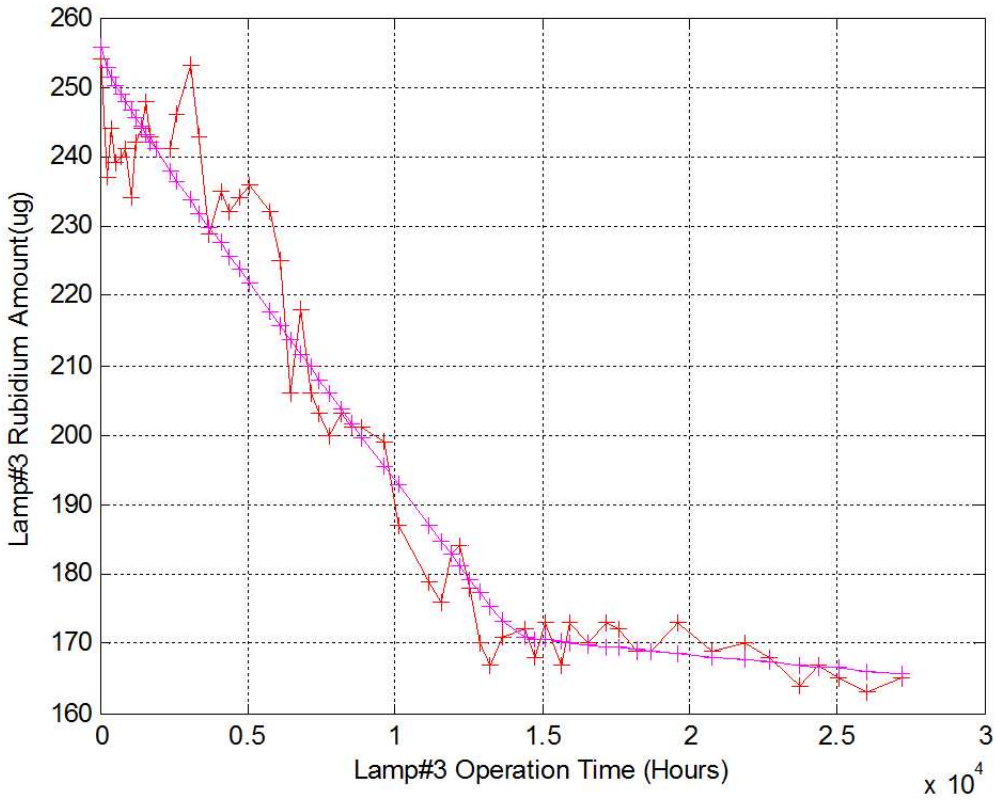
79x62mm (300 x 300 DPI)



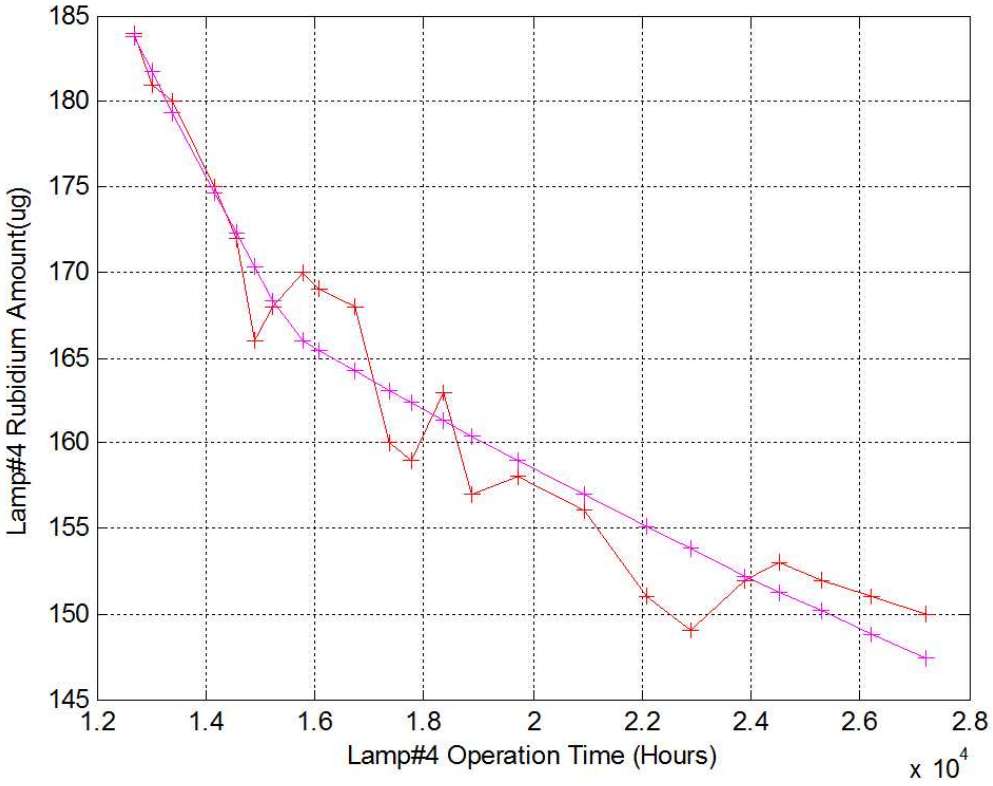
80x63mm (300 x 300 DPI)



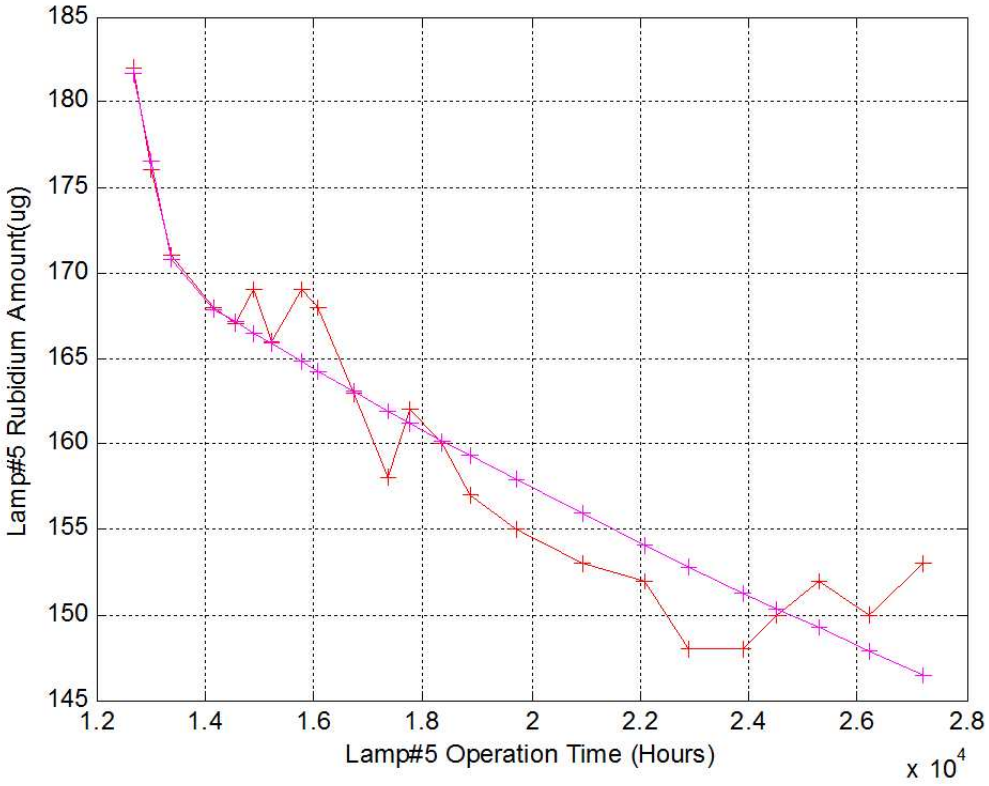
76x61mm (300 x 300 DPI)



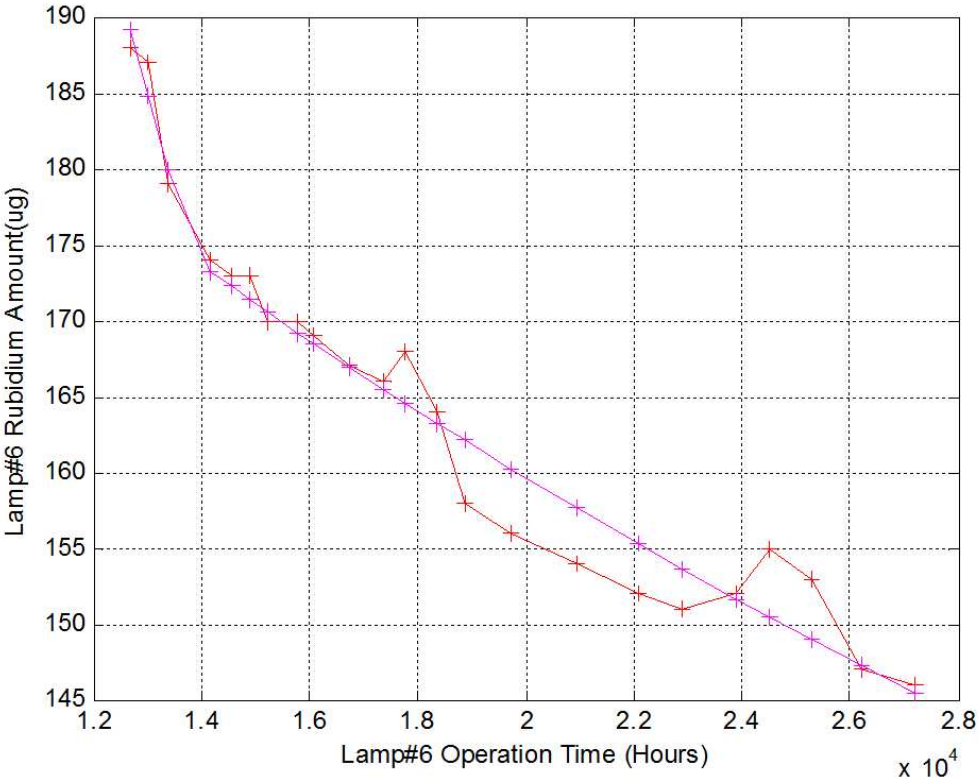
79x63mm (300 x 300 DPI)



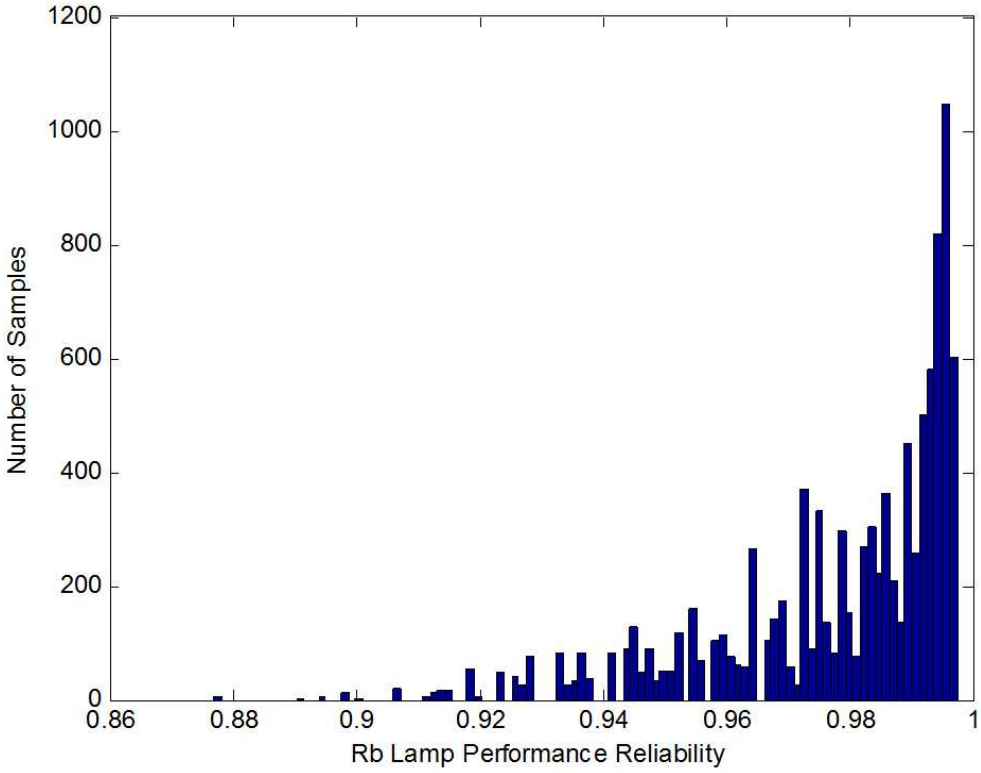
80x62mm (300 x 300 DPI)



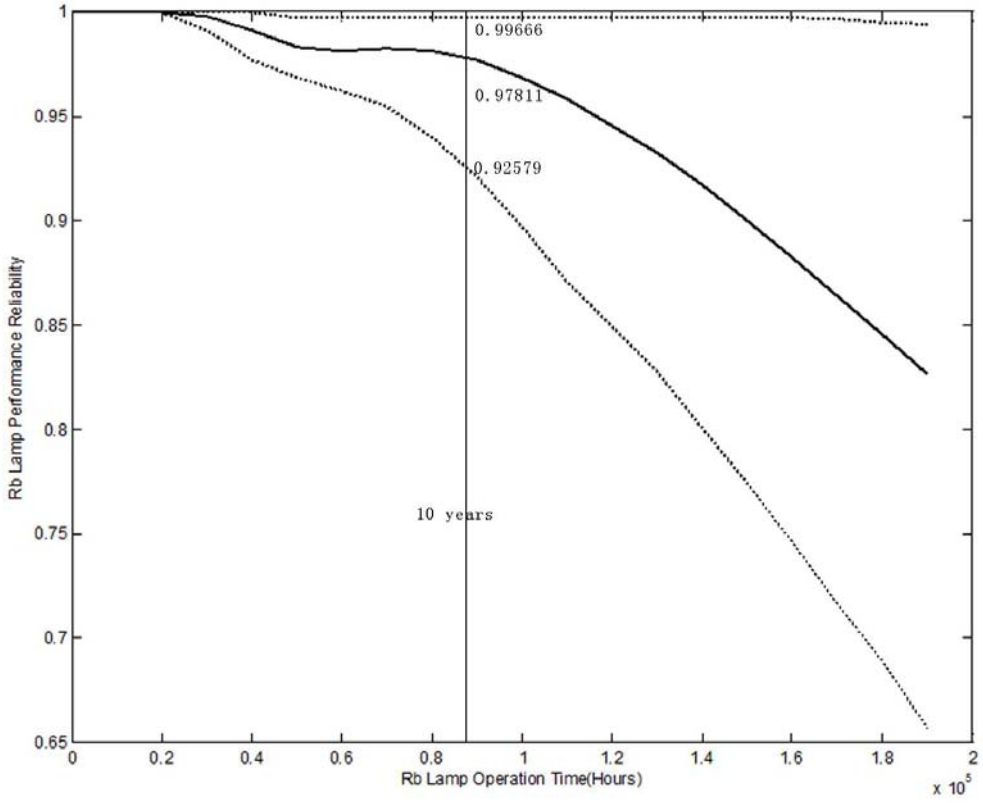
79x62mm (300 x 300 DPI)



79x62mm (300 x 300 DPI)



79x62mm (300 x 300 DPI)



121x101mm (300 x 300 DPI)

Statistical Models for Hot Electron Degradation in Nano-Scaled MOSFET Devices

Suk Joo Bae **Seong-Joon Kim**

Department of Industrial Engineering, Hanyang University, Seoul, Korea

Way Kuo

Department of Electrical & Computer Engineering, The University of Tennessee, Knoxville

Paul H. Kvam

School of Industrial & Systems Engineering, Georgia Institute of Technology

Abstract

In a MOS structure, the generation of hot carrier interface states is a critical feature of the item's reliability. On the nano-scale, there are problems with degradation in transconductance, shift in threshold voltage, and decrease in drain current capability. Quantum mechanics has been used to relate this decrease to degradation and device failure. Although the lifetime and degradation of a device are typically used to characterize its reliability, in this paper we model the distribution of hot-electron activation energies, which has appeal because it exhibits two-point discrete mixture of logistic distributions. The logistic mixture presents computational problems that are addressed in simulation.

Index Terms– EM Algorithm, Logistic Distribution, Maximum Likelihood, Mixture Distribution, Nanotechnology, Reliability.

ACRONYMS

cdf	cumulative distribution function
HCI	hot carrier injection
IC	integrated circuit
MLE	maximum likelihood estimator
$\mathcal{MN}$	multivariate normal
MOS	metal-oxide-semiconductor
MOSFET	metal-oxide-semiconductor field effect transistor
pdf	probability density function

NOTATION

C_i	constant of hot carrier induced degradation model for $i = 1, 2, 3$
$D_{IT}(t_0)$	original interface trap density
$\Delta D_{IT}(t)$	hot carrier activated trap density at time t
$\mathcal{D}(t)$	degradation of a MOSFET device at time t
D_{IT}	interface traps density
E_m	electrical field
$F(\cdot)$	distribution of the hot-electron activation energies
$I(\cdot)$	Fisher information matrix
I_D	drain current
I_{Sub}	substrate current
k	reaction constant
$l(\cdot)$	log-likelihood function
n_b	concentration of Si-H bonds at the interface
n_0	initial concentration of Si-H bonds at the interface
N_T	total concentration of Si bonds
$N_{IT}(t_0)$	initial concentration of interface traps for $t_0 = 0$

$N_{IT}(t)$	concentration of interface traps at time t
$\Delta N_{IT}(t)$	generated interface traps
p	probability of higher activation energies; $p \in (0, 1)$
τ, τ_1, τ_2	lifetime constants
Si^*	Si dangling bond
W	channel width of a device
V_{DD}	power supply voltage
β	coefficient for the I_{Sub}-V_{DD} relationship
μ, σ	parameters of a logistic distribution
μ_i, σ_i	parameters of a mixture logistic distribution for $i = 1, 2$
$\Theta(\Theta_n)$	parameter space (with samples size n)
φ_{IT}	critical energy in electronvolts (eV) for generating an interface trap
φ_0	minimum energy (eV) that an electron must possess to create impact ionization
$\bar{\varphi}_{IT}$, $\bar{\varphi}_{IT,1}$, $\bar{\varphi}_{IT,2}$	mean defect energies
q	elementary charge with the value $1.60218 \times 10^{-19} C$
λ	hot-electron mean-free-path

1 Introduction

The study of reliability has played a vital role in the engineering of products, both large scale and micro scale. In the next decade, it will play an even bigger role for industries in *nanofabrication*, which amounts to designing and manufacturing devices on the nanometer scale; a nanometer (1 nm = one billionth of a meter) is approximately the length of a row of ten hydrogen atoms.

Actually, standard reliability analysis is already essential for the efficient manufacture of nano-devices, but the field of nanotechnology is virtually devoid of results that address reliability issues that are unique to this scale of product. In fact, just as basic physics principles must be rethought at the quantum level, current reliability theories and methods are only partially applicable to systems operating on a nanometer scale. On the molecular level, familiar material properties like conductivity no longer obey laws based on macro scale materials (e.g., Ohm's law). In the same sense, the essential metrics of reliability analysis - material degradation, fatigue, and basic failure mechanisms assume new meaning on the nanometer scale. Sennhauser [25] noted that traditional reliability models may be insufficient due to quantum effects, thermal processes and defect diffusion processes. Experimenters need to consider additional sources of variation such as thermal fluctuations, quantum statistics and Heisenburg uncertainty [3].

There is great potential for reliability improvement if only because current nano-devices are riddled with defects that cause frequent failure problems; the devices are easily damaged by defects that are otherwise harmless to larger micro-devices. A full understanding of the physics and statistics of the defect generation is required in order to investigate the ultimate reliability limitations for nano-devices.

In a MOS structure, for example, the generation of hot carrier interface state is a critical feature of the item's lifetime measurement. Gate current of MOSFETs is made up of electrons injected into the gate oxide by quasi-elastic scattering [13]. However, electrons with high kinetic energies (called "hot carriers") can generate electron-hole pairs near the drain due to impact ionization from atomic-levelled collisions. Those carriers may be injected into the gate oxide and trapped

on defect sites in the oxide. It results in creation of interface states at Si-SiO_2 interface which leads to degradation in transconductance, shift in threshold voltage, and decrease in drain current capability [15]. Understanding the physical mechanisms of HCI will provide meaningful clues for backtracking from observed macro-defects to inferred nano-defects scattered inside the MOSFET devices. This is analogous to reliability problems in which system failure data are used to infer properties about the system's components. In this paper, we investigate physical models of the defects (hot carriers) generation leading to failure based on statistical properties for MOSFETs.

The main results are contained in Section 2. A model for hot electron degradation is achieved via the mixture distribution of hot-electron activation energies. Procedures for statistical inference are outlined in Section 3, and Section 4 contains a discussion of its computation.

2 Physical Models for Hot Carrier Interface State Generation

In a MOS structure, a thin layer of silicon dioxide (SiO_2) forms the insulating layer between the control gate and the conducting channel of transistors used in modern ICs (see Figure 1). As circuits have been made denser to meet the increasing demand for faster logic and memory devices, the dimensions of the transistors have been reduced ("scaled") correspondingly. For example, SiO_2 layer thickness has decreased to 2.0 nm or less, but technology cannot shrink these dimensions indefinitely because thinning down the oxide thickness raises severe technological problems: dielectric thickness variation, penetration of impurities from the highly doped polysilicon gate, reliability and lifetime problems for devices made with the ultrathin oxides, etc [23].

In particular, hot carrier induced degradation in SiO_2 films is perceived as a main potential obstacle for the continued downscaling of MOSFET devices. During device operation, the film is subjected to electrical stress, and electronic defects like hot carriers that limit device lifetimes are more likely to be created for short-channel devices. Generally, silicon-based transistors are

annealed[1] in a hydrogen-rich environment in order to passivate[2] defects at the Si-SiO_2 interface. However, hydrogen (H) is known to play a key role in the HCI degradation of the transistors with smaller geometries. Lyding et al. [15] proposed to replace hydrogen with deuterium during the final wafer sintering process[3] in order to reduce susceptibility to hot electron degradation effects. The details as to how hydrogen degrades a MOSFET device will be illustrated at the following.

2.1 Mechanisms of Hydrogen Release from the Si-SiO_2 Interface

A principal mechanism of MOSFET degradation is the creation of an interface state (or traps) at the Si-SiO_2 interface. The creation is mainly caused by desorption of hydrogens from the passivated dangling bonds at the mismatched Si-SiO_2 interface. This depassivation is activated directly by the hot electrons that exist during transistor switching. The hot electrons near the drain (see Figure 1) in short-channel devices can generate electron-hole pairs via impact ionization[4] [9]. Figure 2 describes the depassivation procedure at Si-SiO_2 interface. First, electrons (e) or holes (h) with high kinetic energies are attracted to the Si-SiO_2 interface, which weaken the Si-H bond until it breaks. As a result, the hydrogen diffuses into the oxide or Si substrate, subsequently creating interface traps with density D_{IT}. The hydrogen (H: atom and H^+: ion) release reaction breaking Si-H bonds is described in the following equation:

$$Si - H + H^+ + e^- \rightarrow Si^* + H_2 \qquad \text{or} \qquad Si - H + H \rightarrow Si^* + H_2, \tag{1}$$

where Si^* represents the Si dangling bond that is an "interface trap".

The hot-carrier-induced trap density D_{IT} is directly proportional to the concentration of Si dangling bonds at the interface. Because the amount of degradation of a MOSFET switching current that leads to device failure is a function of the variation in the interface trap density activated by the hot carriers, the amount of degradation of a MOSFET device can be represented

[1]The annealing is a process that the transistors are heated at sufficiently high temperatures and slowly cooled down

[2]To treat a subject in order to reduce the chemical reactivity

[3]A process of forming a coherent mass by heating without melting

[4]The formation of or separation into ions by heat, electrical discharge, radiation, or chemical reaction

as the concentration of Si dangling bonds at the interface, which can be measured by D_{IT} via charge pumping technique [8].

Electrons need sufficient activation energies to surmount a surface energy barrier to generate interface traps, and the activation energies are directly linked to interface trap defects. Defects do not necessarily have the same activation energy; in fact, there exists a distribution of activation energies for the HCI generation failure mechanism [17]. From a chemistry viewpoint, there are two reasons to expect a distribution of activation energies for MOSFET devices: the variation in the bond energies due to Si-SiO_2 interface disorder, and the possibility of multiple pathways to activation. If the bonding energies are homogeneously distributed at the interface, the activation energy distribution will be of unimodal form, but the distribution will be bimodal if there exist competing mechanisms of interface state formation following multiple pathways to activation [9].

2.2 Activation Energy Distribution of Hot Carrier Induced Defects

Because the activation of hydrogen at the passivated Si-SiO_2 interface is caused by collisions with electrons (or holes) flowing in the channel, it is crucial to identify the energy distribution of these electrons as a function of the number of interface traps over time to evaluate reliability of MOSFET devices.

N_{IT} is proportional to the concentration of the Si dangling bonds at the interface $N_T - n_b$, where N_T is a total concentration of Si bonds which are able to appear as dangling ones if hydrogen leaves the bond. The time dependent trap generation can be described by a simple version of power law [19]:

$$\Delta N_{IT}(t) = N_{IT}(t) - N_{IT}(t_0) = \frac{n_0}{1 + (kt)^{-\alpha}}. \tag{2}$$

The reaction constant k and the power α are values which can be estimated from the experimental data. In terms of the concentration of interface traps, the degradation of a MOSFET device can be approximated by

$$\mathcal{D}(t) = \frac{\Delta D_{IT}(t)}{D_{IT}(t_0)} \simeq \frac{\Delta N_{IT}(t)}{N_{IT}(t_0)} = \frac{1}{1 + (t/\tau)^{-\alpha}}, \tag{3}$$

where $\tau = 1/k$ is a lifetime constant that has units of time. Note that the model (3) is identical to the degradation model for grating decays in optical interconnects derived from Bragg grating theory in Erdogan et al [6]. For the hot carrier degradation mechanism, lifetime determination is based on the observed accelerated degradation of drain voltage. This is because hot carrier degradation is not accelerated by an increase in temperature [10]. The degradation of a device in terms of $\Delta N_{IT}(t)$ can be related to at I_D [13] as

$$\Delta N_{IT}(t) = C_1 \left[t \frac{I_D}{W} \exp\left(-\frac{\varphi_{IT}}{q\lambda E_m} \right) \right]^n . \tag{4}$$

Introducing an easily measurable I_{Sub} to monitor the device degradation, E_m can be represented with the multiplication factor

$$\frac{I_{Sub}}{I_D} = C_2 \exp\left(-\frac{\varphi_0}{q\lambda E_m} \right). \tag{5}$$

A lifetime is defined as the time to reach a fixed number of interface traps. By combining (4) and (5), the lifetime for a specific MOSFET device, t_f is

$$\frac{t_f I_D}{W} \propto \left[\frac{I_{Sub}}{I_D} \right]^{-\varphi_{IT}/\varphi_0} \qquad \text{or} \qquad t_f \propto I_{Sub}^{-\varphi_{IT}/\varphi_0}. \tag{6}$$

The substrate current is a function of the power supply voltage as

$$I_{Sub} \propto \exp\left(-\frac{\beta}{V_{DD}} \right). \tag{7}$$

Combining (6) and (7), the lifetime of the device can be written as

$$t_f = C_3 \exp\left(\frac{\varphi_{IT}\beta}{\varphi_0 V_{DD}} \right). \tag{8}$$

Here, t_f also connotes the time needed to activate Si-H bonds with activation energy less than φ_{IT}. Following the approach in [9], we replace $C_3 \exp(\bar{\varphi}_{IT}\beta/(\varphi_0 V_{DD}))$ by τ in (3) which can be estimated by fitting integrated coupling constant (ICC) to decay time [6]. Finally the degradation of a MOSFET device can be approximated by using the following distribution on the hot-electron activation energies:

$$F(\varphi_{IT}) = \left[1 + \exp\left(-\frac{\varphi_{IT} - \bar{\varphi}_{IT}}{\sigma} \right) \right]^{-1}, \tag{9}$$

where $\sigma = (\varphi_0 V_{DD})/(\alpha\beta)$. Note that the degradation model $\mathcal{D}(t)$ in (3) that is represented as the proportion of activated defects before time t is equivalent to the probability that activated defects have activation energy less than or equal to $\bar{\varphi}_{IT}$.

Although the lifetime and degradation of a device are typically used to characterize its reliability, in this case the distribution of hot-electron activation energies has a unique appeal because it has a common logistic structure. The logistic distribution, derived from disorder-induced variations in the *Si-H* activation energies, is identical to a Fermi-derivative distribution of the energies of electronic states [5]. Via accelerated degradation test of a MOSFET device, we can measure HCI degradation over time and then extract the distribution of hot-electron activation energies. Figure 3-(a) shows how the degradation model caused by interface traps (or defects) varies as a function of power supply voltage(V_{DD}) of a MOSFET device for fixed values of β and φ_0, along with the distributions of defect activation energies in Figure 3-(b). The parameter values in the figure are simulated from experimental observations of short-time-tests for 180 nm MOSFET devices in Haggag et al. [9].

2.3 Bimodal Distribution of Activation Energies

Existence of multiple paths and competing mechanisms for the release of hydrogen yields inhomogeneous activation energy distributions. Figure 4 shows the energy level of hydrogen release to different activation pathways (the hydrogen may be attracted to *Si* or SiO_2 or the *Si*-SiO_2 interface). Through atomic simulations based on density functional theory, Tuttle et al. [28] showed that the activation energy of hydrogen is distributed around 3.5 eV if the hydrogen desorbs into the SiO_2, but below 3 eV if the final hydrogen state is closer to the silicon bulk. As a result, the time-dependent HCI degradation model is a *mixture* of the model (3) [11]:

$$\mathcal{D}(t) = \frac{\Delta D_{IT}(t)}{D_{IT}(t_0)} = \frac{p}{1 + (t/\tau_1)^{-\alpha_1}} + \frac{1-p}{1 + (t/\tau_2)^{-\alpha_2}}. \tag{10}$$

By letting $\tau_1 = C_{3,1} \exp(\bar{\varphi}_{IT,1}\beta_1/(\varphi_0 V_{DD}))$ and $\tau_2 = C_{3,2} \exp(\bar{\varphi}_{IT,2}\beta_2/(\varphi_0 V_{DD}))$ in (10), the degradation model of a MOSFET device can be represented through φ_{IT} as a mixture of logistic distri-

butions:

$$F(\varphi_{IT}) = p \cdot \left[1 + \exp\left(-\frac{\varphi_{IT} - \bar{\varphi}_{IT,1}}{\sigma_1}\right)\right]^{-1} + (1-p) \cdot \left[1 + \exp\left(-\frac{\varphi_{IT} - \bar{\varphi}_{IT,2}}{\sigma_2}\right)\right]^{-1}, \qquad (11)$$

where $\sigma_1 = (\varphi_0 V_{DD})/(\alpha_1\beta_1)$ and $\sigma_2 = (\varphi_0 V_{DD})/(\alpha_2\beta_2)$.

Tuttle et al. [28] experimentally observed a higher mean energy $\bar{\varphi}_{IT,1} \approx 3.5$ eV as well as a lower mean energy $\bar{\varphi}_{IT,2} \approx 2.9$ eV. The higher energy band comes from "single collisions" with higher energetic electrons and a consequent release of the hydrogen through a higher energy path in the MOSFET. On the other hand, the lower energy band comes from "multiple collisions" with lower energetic electrons and a consequent release of the hydrogen through a lower energy path in the MOSFET [9]. Figures 5-(a) and 5-(b) display the mixture of time-dependent HCI degradation model (10) and the mixture distribution of defect activation energies, respectively, at varying p values with $\bar{\varphi}_{IT,1} = 3.5$ eV, $\bar{\varphi}_{IT,2} = 2.9$ eV, and $V_{DD} = 3.0V$.

3 Parameter Estimation

We showed the degradation of a MOSFET device can be approximated by a logistic or a mixture of logistic distribution on the hot-electron activation energies through step-by-step procedure at previous section. The MOSFET lifetime is defined as the time that the actual degradation path reaches the prespecified degradation threshold level. In terms of hot-electron activation energies, the lifetime is equivalent to the activation of a sufficient number of hydrogen with activation energy less then or equal to $\bar{\varphi}_{IT}$ [11]. In this section, we outline the procedures for statistical inference for different characteristics of the MOSFET lifetime. Using the measurement of hot-electron activation energies, we rely on the method of maximum likelihood to estimate logistic model parameters, or more precisely, parameters for the logistic mixture distribution. While the inference for the logistic distribution is straightforward, there are important issues in dealing with estimation for the mixture distribution.

3.1 Logistic Mixture Distribution

The cdf of the random variable X having the logistic distribution is given by

$$F(x;\boldsymbol{\theta}) = \frac{1}{\left\{1+\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}}, \qquad -\infty < x < \infty, \tag{12}$$

for $\boldsymbol{\theta} = (\mu, \sigma)^T$, where μ and σ are location and scale parameters. The corresponding pdf is

$$f(x;\boldsymbol{\theta}) = \frac{1}{\sigma}\frac{\exp\left(-\frac{x-\mu}{\sigma}\right)}{\left\{1+\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}^2} = \frac{1}{\sigma}\left\{F(x)[1-F(x)]\right\}. \tag{13}$$

The pdf of the logistic distribution is symmetric and bell-shaped like that of the normal distribution. Since the logistic distribution has slightly longer tails, it would require an extremely large number of observations to accurately assess whether data come from a normal or logistic distribution. The logistic random variable X has mean $E[X] = \mu$, variance $\mathrm{Var}(X) = (\pi^2\sigma^2)/3$ and coefficient of variation $\pi\sigma/(\mu\sqrt{3})$. The log-likelihood function for a sample of size n from the logistic distribution is given by

$$l(x_1,\ldots,x_n;\boldsymbol{\theta}) = \sum_{i=1}^{n}\log f(x_i;\boldsymbol{\theta}) = -n\log\sigma - \sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right) - 2\sum_{i=1}^{n}\log\left\{1+\exp\left(-\frac{x_i-\mu}{\sigma}\right)\right\},$$

and the MLEs, $(\hat{\mu}, \hat{\sigma})$ of the parameters (μ, σ) satisfy the following likelihood equations:

$$\begin{aligned} \sum_{i=1}^{n}\left\{1+\exp\left(-\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right)\right\}^{-1} &= \frac{n}{2}, \\ -\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right) + \sum_{i=1}^{n}\left(\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right)\left\{1+\exp\left(-\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right)\right\}^{-1} &= \frac{n}{2}. \end{aligned} \tag{14}$$

Taking advantage of the similarity in shape between the logistic and normal distributions, initial values of $\hat{\mu}$ and $\hat{\sigma}$ might be taken as $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$, and $\sqrt{n^{-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}$, respectively. Then solutions could be improved by applying the Newton-Raphson method. When both μ and σ are unknown, the Newton-Raphson method converges quickly to the solutions $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma})^T$. Since the logistic-likelihood function is quasi-concave, the solutions are unique for distinct values of x_i. [1].

The MLEs $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma})^T$, as consistent roots of the likelihood equations (14), satisfy

$$\left(\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)\right) \xrightarrow{L} \mathcal{MN}(\mathbf{0}, I^{-1}(\boldsymbol{\theta}_0)), \tag{15}$$

where $\boldsymbol{\theta}_0$ is the true value of $\boldsymbol{\theta}$ and the Fisher information $I(\boldsymbol{\theta}_0)$ is given by

$$I(\boldsymbol{\theta}_0) = -E \begin{bmatrix} \frac{\partial^2 l}{\partial^2 \mu^2} & \frac{\partial^2 l}{\partial \mu \partial \sigma} \\ \frac{\partial^2 l}{\partial \mu \partial \sigma} & \frac{\partial^2 l}{\partial^2 \sigma^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{3\sigma^2} & 0 \\ 0 & \frac{3+\pi^2}{9\sigma^2} \end{bmatrix}. \tag{16}$$

It is common in practice to estimate the inverse of the covariance matrix of the MLE by the observed information matrix $I(\hat{\boldsymbol{\theta}})$ rather than the expected information matrix $I(\boldsymbol{\theta}_0)$ evaluated at $\boldsymbol{\theta}_0 = \hat{\boldsymbol{\theta}}$. In general, the observed information matrix is more convenient to use than the expected information matrix as it does not require an expectation to be taken. However as shown in (16), the expectations are trivial in the logistic case and we can easily derive the covariance matrix of the MLE from the expected information matrix. When only μ is unknown, a MLE $\hat{\mu}$ can be uniquely determined by replacing $\hat{\sigma}$ with known σ value in the likelihood equation. Alternatively, we can use an estimator

$$\zeta_n = \bar{X}_n - \frac{l'(\bar{X}_n)}{l''(\bar{X}_n)}, \tag{17}$$

instead of $\hat{\mu}$ because they have the same asymptotic distribution. Here $\bar{X}_n$ is the average of n samples as a $\sqrt{n}$-consistent estimator (denoted as $\tilde{\mu}$) and

$$l'(\bar{X}_n) = \frac{\partial}{\partial \mu} l(\mathbf{x}; \boldsymbol{\theta})\big|_{\mu=\bar{X}_n} = n - 2\sum_{i=1}^{n} \frac{\exp\left(\frac{x_i - \bar{x}_n}{\sigma}\right)}{1 + \exp\left(\frac{x_i - \bar{x}_n}{\sigma}\right)},$$

$$l''(\bar{X}_n) = \frac{\partial^2}{\partial \mu^2} l(\mathbf{x}; \boldsymbol{\theta})\big|_{\mu=\bar{X}_n} = -2\sum_{i=1}^{n} \frac{\exp\left(\frac{x_i - \bar{x}_n}{\sigma}\right)}{1 + \exp\left(\frac{x_i - \bar{x}_n}{\sigma}\right)^2}.$$

Theorem 3.1 *Let $\hat{\mu}$ be the MLE of μ, and let ζ_n be given by (17), then $\sqrt{n}(\zeta_n - \hat{\mu}) \to 0$ as $n \to \infty$.*

The proof is listed in the appendix. When only μ is known, the MLE $\hat{\sigma}$ can be uniquely determined by replacing $\hat{\mu}$ with known μ value in the likelihood equations. With moment estimator $\tilde{\sigma} = \{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2/n\}^{1/2}$, which is a $\sqrt{n}$-consistent estimator of σ, ,

$$\eta_n = \tilde{\sigma} - \frac{l'(\tilde{\sigma})}{l''(\tilde{\sigma})} \tag{18}$$

has the same asymptotic distribution as the MLE $\hat{\sigma}$ (from the theorem) with $\tilde{\sigma}$ instead of $\bar{X}_n$. Here,

$$l'(\tilde{\sigma}) = \frac{\partial}{\partial \sigma} l(\mathbf{x}; \boldsymbol{\theta})\big|_{\sigma=\tilde{\sigma}} \qquad \text{and} \qquad l''(\tilde{\sigma}) = \frac{\partial^2}{\partial \sigma^2} l(\mathbf{x}; \boldsymbol{\theta})\big|_{\sigma=\tilde{\sigma}}.$$

The degradation model of a MOSFET device can represented in (11) is a two-point discrete mixture of logistic distributions. For a random variable X generated from this mixture of logistic distributions, then in terms of $\boldsymbol{\theta}_1 = (\mu_1, \sigma_1)^T$, $\boldsymbol{\theta}_2 = (\mu_2, \sigma_2)^T$ and $\boldsymbol{\Psi} = (p, (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T))^T$, X has pdf

$$f(x; \boldsymbol{\Psi}) = p \cdot f_1(x; \boldsymbol{\theta}_1) + (1 - p) \cdot f_2(x; \boldsymbol{\theta}_2), \tag{19}$$

and corresponding cdf

$$F(x; \boldsymbol{\Psi}) = p \cdot F_1(x; \boldsymbol{\theta}_1) + (1 - p) \cdot F_2(x; \boldsymbol{\theta}_2), \tag{20}$$

where $F_j(x; \boldsymbol{\theta}_j)$ and $f_j(x; \boldsymbol{\theta}_j)$ are from (12) and (13), respectively, with parameters μ_j and σ_j for $j = 1, 2$.

4 Solving the MLE

Several methods have been proposed to estimate the parameter $\boldsymbol{\Psi}$. The MLE, based on maximizing the log-likelihood function

$$l(x_1, \ldots, x_n; \boldsymbol{\Psi}) = \sum_{i=1}^{n} \log \left\{ p \cdot f_1(x_i; \boldsymbol{\theta}_1) + (1 - p) \cdot f_2(x_i; \boldsymbol{\theta}_2) \right\},$$

possesses a number of desirable statistical properties. The computation of the MLE of $\boldsymbol{\Psi}$ requires solving the likelihood equation $\partial l(x_1, \ldots, x_n; \boldsymbol{\Psi}) / \partial \boldsymbol{\Psi} = \mathbf{0}$. The resulting MLE of $\boldsymbol{\Psi}$, $\hat{\boldsymbol{\Psi}}$ satisfies

$$\hat{p} = \sum_{i=1}^{n} \kappa(x_i; \hat{\boldsymbol{\Psi}}) / n \qquad \text{and} \qquad \sum_{j=1}^{2} \sum_{i=1}^{n} \kappa(x_i; \hat{\boldsymbol{\Psi}}) \partial \log f_j(x_i; \hat{\boldsymbol{\theta}}_j) / \partial \boldsymbol{\theta} = \mathbf{0}, \tag{21}$$

where $\kappa(x_i; \hat{\boldsymbol{\Psi}}) = \hat{p} f_1(x_i; \hat{\boldsymbol{\theta}}_1) / [\hat{p} f_1(x_i; \hat{\boldsymbol{\theta}}_1) + (1 - \hat{p}) f_2(x_i; \hat{\boldsymbol{\theta}}_2)]$ is the posterior probability that x_i belongs to the first component of the mixture.

The EM algorithm (see McLachlan and Krishnan [16], for example) can be used to find the MLE for mixtures by solving the likelihood equations (21) iteratively. To apply the EM method, we imagine each observation from the mixture distribution comes with an indicator variable that tells us which of the two logistic distributions the observation was generated. In this case, such an indicator is treated as a missing value.

Starting from an arbitrary initial guess, the algorithm operates in two repeated steps. The E-step estimates missing values as they appear in the log-likelihood, then the M-step finds a local optimum to the likelihood using the estimated data in place of what was missing. With the estimated indicator functions, the MLE is solved more simply using (14) in the M-step (as discussed in the last section). Convergence properties for finite mixture models are discussed in Tanaka and Takemura [22]. In this case, because of the heavy tails in the logistic density and the large variances in (16), the convergence can be slow, depending on how good the initial guess is.

In Table 1, we present summary results of the simulation comparing the performance of different MLE/EM algorithms for mixtures of logistic components. Comparisons between varying degrees of separation in the mixture distributions (complete separation, moderate overlap, and large overlap) are illustrated in Figures 6 (a) - (c), respectively. In these simulations, the mixing proportion p takes on the values .20, .50, and .80. For a given mixture, the component distributions differ from each other only by location and scale differences. For each set of parameter configurations, samples of size $n = 1,000$ were generated from the corresponding mixture of logistic distributions.

Data generation and parameter estimation for logistic mixture were executed using the *mle* program [12], and for maximizing the likelihood function in the EM algorithm, four different methods were used; simplex, direct, conjugate gradient, and simulated annealing (see [20] for details of the four methods).

Each of the methods has strengths and weakness for different types of functions. In the case of the logistic likelihood, they have peculiar differences; Table 1 shows a summary of the simulation comparing the performance of MLE/EM algorithms for mixtures of logistic components. The computation stopped when the relative errors ϵ, of all 5 parameters, $\mathbf{\Psi} \equiv (p, \mu_1, \sigma_1, \mu_2, \sigma_2)'$ reached $10^{-4} \cdot \epsilon \equiv [\mathbf{\Psi}^{(h+1)} - \mathbf{\Psi}^{(h)}]/\mathbf{\Psi}^{(h)}$. The symbol $(*)$ represents result tends to a pathological solution and the relative bias, $|y - \hat{y}|/y$ is calculated over every parameter in parentheses. The direct and simulated annealing methods provide better results than the other two methods for this parameter estimation. However, simulated annealing requires longer process time for the annealing function.

In general, the estimation precision increases when the mixing proportion is neutral ($p = 0.5$).

One way of obtaining standard errors of the estimates of the parameters in a mixture model is to approximate the covariance matrix of $\hat{\mathbf{\Psi}}$ by the inverse of the observed information matrix. For mixture models, the sample size has to be very large to guarantee the asymptotic theory of maximum likelihood, hence a resampling approach such as bootstrapping method can be considered to construct standard errors of the estimates of the parameters. Standard error estimation of $\hat{\mathbf{\Psi}}$ can implemented via a bootstrap procedure; Chapter 13 of [18] for an analogous resampling approach.

5 Discussion

Statistical models based on known physical principles (e.g., the power-rule model, the Arrhenius rule, Eyring Model, etc.) have provided strong methods for parametric inference for various testing problems in manufacturing. Nano-manufacturing will provide more twists to these traditional models due to the nature of nano-defects and Heisenberg uncertainty. This paper provides basic physical modeling for MOSFET devices based on the nano-level degradation that takes place at defect sites in the MOSFET gate oxide. The distribution of hot-electron activation energies proves to be more accessible than analogous measures of degradation or lifetime, and is derived as a logistic mixture distribution using physical principles on the nanoscale. Although the inference problem is ridden with computational challenges, the derivation of MLEs is straightforward using the EM algorithm.

Appendix

Proof of Theorem : For the true value of μ (denoted as μ_0), Taylor expansion of $l'(\hat{\mu})$ about $l'(\mu_0)$ is

$$l'(\hat{\mu}) = l'(\mu_0) + (\hat{\mu} - \mu_0)l''(\mu_0) + \frac{1}{2}(\hat{\mu} - \mu_0)^2 l'''(\mu^*),$$

where μ^* lies between μ_0 and $\hat{\mu}$. $l'''(\cdot)$ exists for the three-times differentiable logistic likelihood function with respect to μ, and it is bounded. Since the left side is zero for the MLE $\hat{\mu}$, we have

$$\sqrt{n}(\hat{\mu}-\mu_0) = -\frac{\sqrt{n}l'(\mu_0)}{l''(\mu_0)} + R_n,$$

where

$$R_n = -\frac{\sqrt{n}l'(\mu_0)}{l''(\mu_0)}\left[\frac{1}{1+\frac{n}{l''(\mu_0)}\frac{1}{2n}(\hat{\mu}-\mu_0)l'''(\mu^*)} - 1\right].$$

Analogously, Taylor expansion of $l'(\tilde{\mu})$ about $l'(\mu_0)$ can be represented as

$$l'(\tilde{\mu}) = l'(\mu_0) + (\tilde{\mu}-\mu_0)l''(\mu_0) + \frac{1}{2}(\tilde{\mu}-\mu_0)^2 l'''(\mu^*).$$

It follows from (15) that

$$\zeta_n = \tilde{\mu} - \frac{1}{l''(\tilde{\mu})}\left[l'(\mu_0) + (\tilde{\mu}-\mu_0)l''(\mu_0) + \frac{1}{2}(\tilde{\mu}-\mu_0)^2 l'''(\mu^*)\right],$$

and by using the expansion $l''(\tilde{\mu}) = l''(\mu_0) + (\tilde{\mu}-\mu_0)l'''(\mu^*)$,

$$\sqrt{n}(\zeta_n-\mu_0) = -\frac{\sqrt{n}l'(\mu_0)}{l''(\mu_0)}\left(1+\frac{(\tilde{\mu}-\mu_0)l'''(\mu^*)}{l''(\mu_0)}\right)^{-1} + R_n',$$

where

$$R_n' = \sqrt{n}(\hat{\mu}-\mu_0)\left[1 - \frac{l''(\mu_0)}{l''(\tilde{\mu})} - \frac{1}{2}(\tilde{\mu}-\mu_0)\frac{l'''(\mu^*)}{l''(\tilde{\mu})}\right].$$

As $n \to \infty$, R_n and R_n' tend to 0 in probability since $\hat{\mu} \to \mu_0$, $\frac{l''(\mu_0)}{l''(\tilde{\mu})} \to 1$, and $\tilde{\mu} \to \mu_0$, and additionally $\left(1+\frac{(\tilde{\mu}-\mu_0)l'''(\mu^*)}{l''(\mu_0)}\right) \to 1$.

References

[1] Antle, C. E., Klimko, L., and Harkness, W. (1970) "Confidence Intervals for the Parameters of the Logistic Distribution", *Biometrika*, Vol. 57, No. 2, 397–402.

[2] Balakrishnan, N. (1992) *Handbook of the Logistic Distribution*, Dekker, New York.

[3] Birge, R. R., Lawrence, A. F., and Tallent, J. R. (1991) "Quantum effects, thermal statistics and reliability of nanoscale molecular and semiconductor devices", *Nanotechnology*, Vol. 2, 73–87.

[4] DeCani, J. S., and Stine, R. A. (1986), "A Note on Deriving the Information Matrix for a Logistic Distribution", *The American Statistician*, Vol. 40, No. 3, 220–222.

[5] Devine, R. A. B., Autran, J. L., Warren, W. L., Vanheusdan, K. L., and Rostaing, J. C. (1997), "Interfacial Hardness Enhancement in Deuterium Annealed 0.25 μm Channel Metal Oxide Semiconductor Transistors", *Applied Physics Letters*, Vol. 70, No. 22, 2999–3001.

[6] Erdogan, T., Mizrahi, V., Lemaire, P. J., and Monroe, D. (1994), "Decay of Ultraviolet-Induced Fiber Bragg Gratings", *Journal of Applied Physics*, Vol. 76, No. 1, 73–80.

[7] Feng, Z. D., and McCulloch, C. E. (1996), "Using Bootstrap Likelihood Ratios in Finite Mixture Models", *Journal of the Royal Statistical Society B*, Vol. 58, No. 3, 609–617.

[8] Groeseneken, G., Maes, H. E., Beltran, N., and DE Keersmaecker, R. F. (1984), "A Reliable Approach to Charge- Pumping Measurements in MOS Transistors", *IEEE Transactions on Electron Devices*, Vol. 31, No. 1, 42–53.

[9] Haggag, A., McMahon, W., Hess, K., Cheng, K., Lee, J., and Lyding, J. W., (2001), "High-Performance Chip Reliability from Short-Time-Tests: Statistical Models for Optical Interconnect and Deep-Submicron Transistor Failures",*International Reliability Physics Symposium*, Orlando, FL.

[10] Heremans, P., Van den Bosch, G., Bellens, R., Groeseneken, G., Maes, H. E. (1990), "Temperature Dependence of the Channel Hot-Carrier Degradtion of n-Channel MOSFETs", *IEEE Transactions on Electron Devices.*, Vol. 37, No. 4., 980–993.

[11] Hess, K., Haggag, A., McMahon, W., Cheng, K., Cheng, K., Lee, J., and Lyding, J. W. (2001), "The Physics of Determining Chip Reliability", *IEEE Circuits Devices Magazine*, Vol. 17, No. 3, 33–38.

[12] Holman, D. J. (2003), *A Programming Language for Buliding Likelihood Models*, User's manual, Vol. 1, Ver. 2.1.

[13] Hu, C., Tam, S. C., Hsu, F., Ko, P., Chan, T., and Terrill, K. W. (1985), "Hot-Electron-Induced MOSFET Degradation-Model, Monitor, and Improvement', *IEEE Journal of Solid-State Circuits*, Vol. 20, No. 1, 295–305.

[14] Lehmann, E. L., and Casella, G. (1998) *Theory of Point Estimation*, Springer, New York.

[15] Lyding, J. W., Hess, K., and Kizilyalli, I. C. (1996), "Reduction of Hot Electron Degradation in Metal Oxide Semiconductor Transistors by Deuterium Processing", *Applied Physics Letters*, Vol 68. No. 18, 2526–2528.

[16] McLachlan, G. J., and Krishnan, T. (1997), *The EM Algorithm and Extensions*, Wiley, New York.

[17] McMahon, W., Haggag, A., and Hess, K. (2003) "Reliability Scaling Issues for Nanoscale Devices", *IEEE Transactions on Nanotechnology*, Vol. 2, No. 1, 33–38.

[18] Meeker, W. Q. and Escobar, L. A. (1998), *Statistical Methods for Reliability Data*, Wiley, New York.

[19] Penzin, O., Haggag, A., McMahon, W., Lyumkis, E., and Hess, K. (2003), "MOSFET Degradation Kinetics and Its Simulation", *IEEE Transactions on Electron Devices*, Vol. 50, No. 6, 1445–1450.

[20] Press, W. H., Flannery B. P., Teukolsky S. A., and Vetterling W. T. (1989), *Numerical Recipes in Pascal: The Art of Scientific Programming*, Cambridge University Press, Cambridge.

[21] Quandt, R. E., and Ramsey, J. B. (1978), "Estimating Mixtures of Normal Distributions and Switching Regressions", *Journal of American Statistical Association*, Vol. 73, 730-738.

[22] Tanaka, K., and Takemura, A. (2003), "Strong Consistency of MLE for Finite Mixtures of Location-Scale Distributions When the Scale Parameters are Exponentially Small", *Mathematical Engineering Technical Report*, University of Tokyo.

[23] Waser, R., (2003) *Nanoelectronics and Information Technology: Advanced Electronic Materials and Novel Devices.*, Wiley, Weinheim.

[24] Sennhauser, U., Reiner, J., and Nellen, P. M. (2004), "Nanoreliability", *13th NID Workshop*, Athens, Greece.

[25] Sennhauser, U. (2004), "Reliability of Nanostructed Materials and Devices", Project No. 5202.1, Swiss Federal Laboratories for Materials Testing and Research.

[26] Shukla, S. K., Norman, G., Parker, D., and Kwiatkowska, M. (2005), "Evaluating the Reliability of Defect-Tolerant Architectures for Nanotechnology with Probabilistic Model Checking", Submitted for Publication in *IEEE Transactions on Computer Aided Design* (under 2nd revision).

[27] Stathis, J. H. and DiMaria, D. J. (1998), "Reliability Projection for Ultra-Thin Oxides at Low Voltage", *IEDS Technical Digest*, Vol. 6, 167–170.

[28] Tuttle, B., and Van d Walle, C. G. (1999), "Structure, Energetics, and Vibrational Properties of Si-H Bond Dissociation in Silicon", *Physical Review B-Condensed Matter and Materials Physics*, Vol. 39, 12884–12889.

[29] Vallett, D. P. (2002) "Failure Analysis Requirements for Nanoelectronics", *IEEE Transactions on Nanotechnology*, Vol. 1, No. 3, 117–121.

Authors

Suk Joo Bae is an Assistant Professor in the Department of Industrial Engineering at Hanyang University, Seoul, Korea. He received his Ph.D. from the School of Industrial and Systems Engineering at the Georgia Institute of Technology in 2003. He worked as a reliability engineer at the SDI, Korea, from 1996 to 1999. His research interests are centered on reliability evaluation of light displays and nano-devices via accelerated life and degradation testing, statistical robust parameter

design, and process control for large-volume on-line processing data. He is a member of INFORMS, ASA, and IMS.

Seong-Joon Kim is currently a Graduate Student in the Department of Industrial Engineering at Hanyang University, Seoul, Korea. His research interests are centered on reliability evaluation of nano-devices, spatial modeling of point process.

Way Kuo is the University Distinguished Professor and Dean of Engineering at the University of Tennessee. He was the Wisenbaker Chair of Engineering and head of the industrial engineering department at Texas A&M. He is an elected member of the National Academy of Engineering. He received his B.S. in nuclear engineering from National Tsing Hua University in Taiwan and his Ph.D in industrial engineering from Kansas State University. He is a fellow of IIE, IEEE, and ASQ, and an Academician of International Academy for Quality. He has chaired the Council of Industrial Engineering Academic Department Heads and the IIE Council of Fellows.

Paul H. Kvam is a Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He received his Ph.D. from the University of California, Davis in 1991. His research interests include reliability evaluation in engineering applications, including accelerated life testing, degradation testing, analysis of dependent systems, and nonparametric inference. He is a fellow of the ASA.

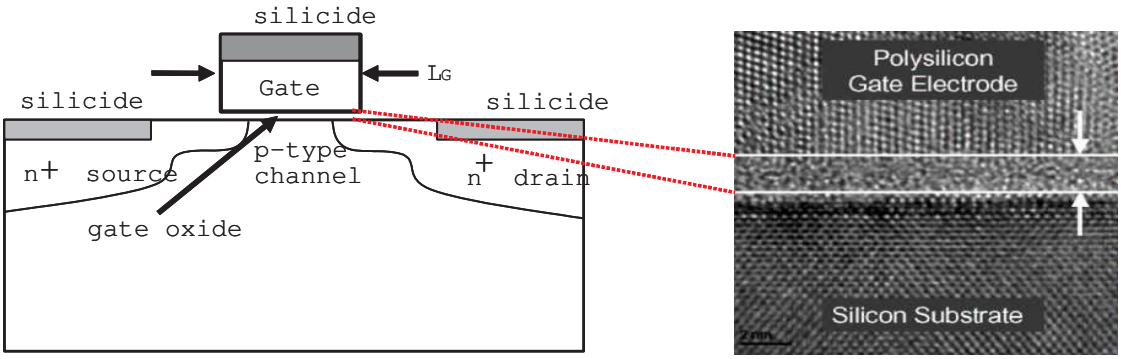


Figure 1: Basic structure of a planar MOSFET: For 90 nm generation gate oxide, the thickness of silicon oxide (SiO_2) is less than 2 nm.

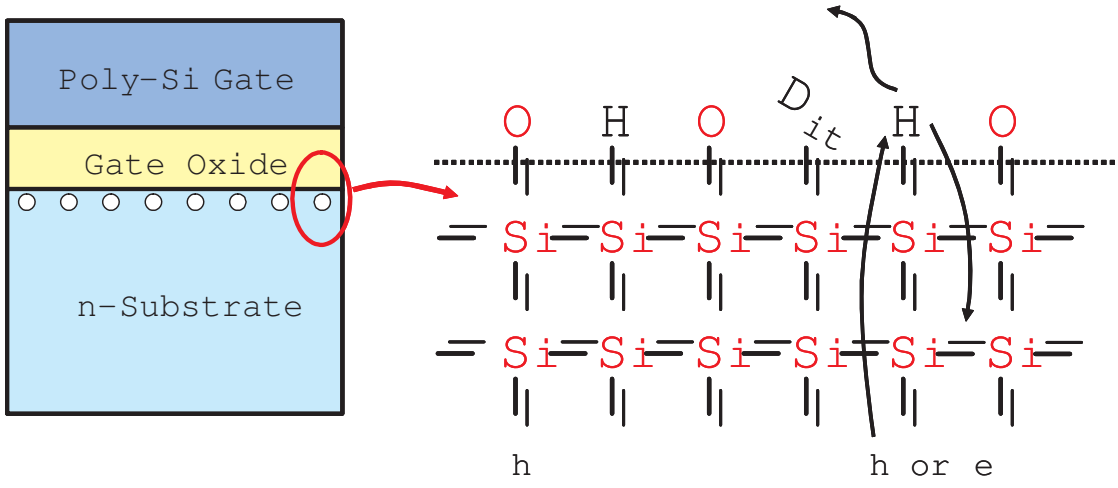
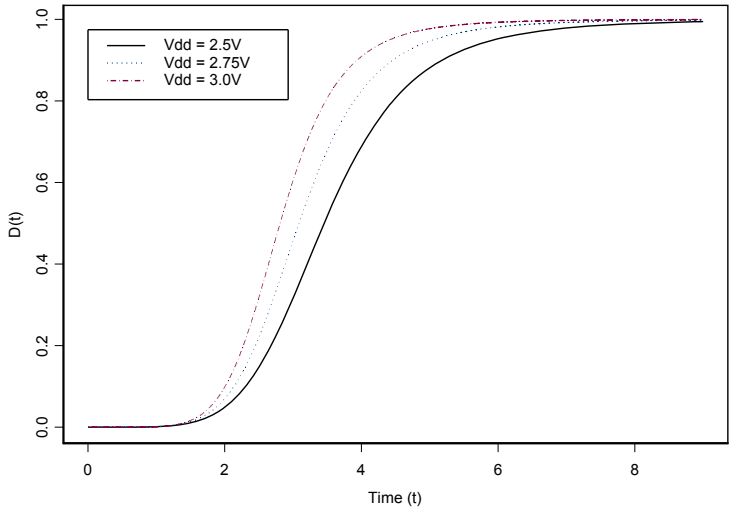
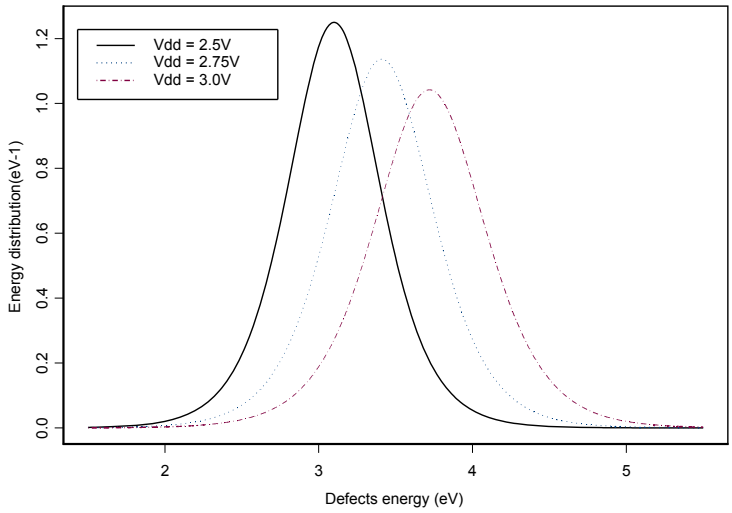


Figure 2: Desorption procedure of hydrogens at Si-SiO_2 interface.



(a)



(b)

Figure 3: (a)-Defects degradation model (3); (b)-Distribution of activation energies for interface traps.

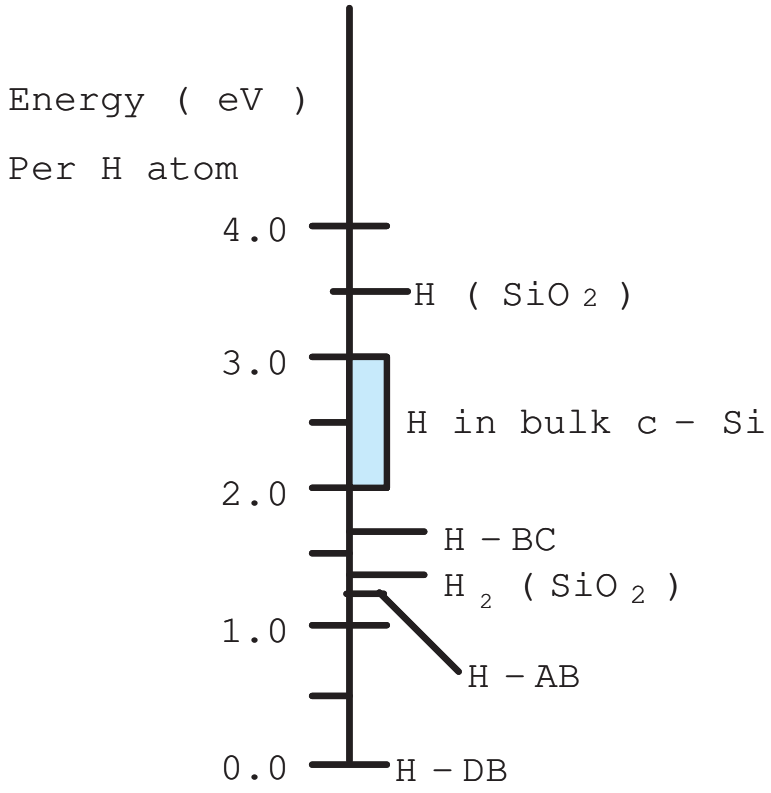
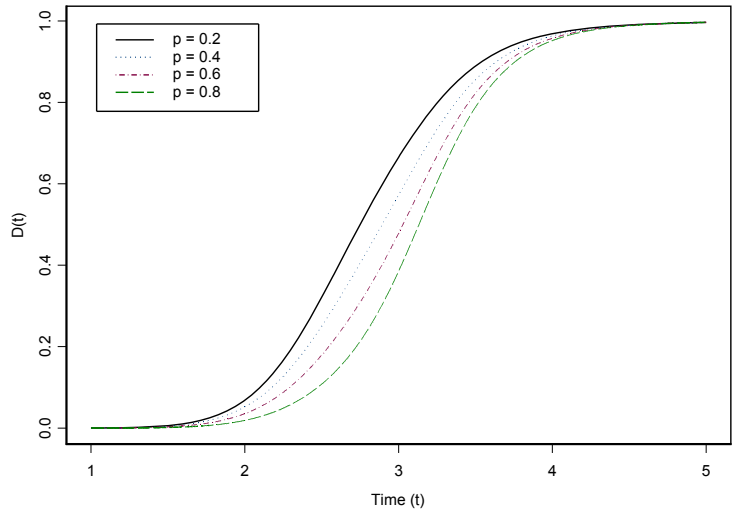
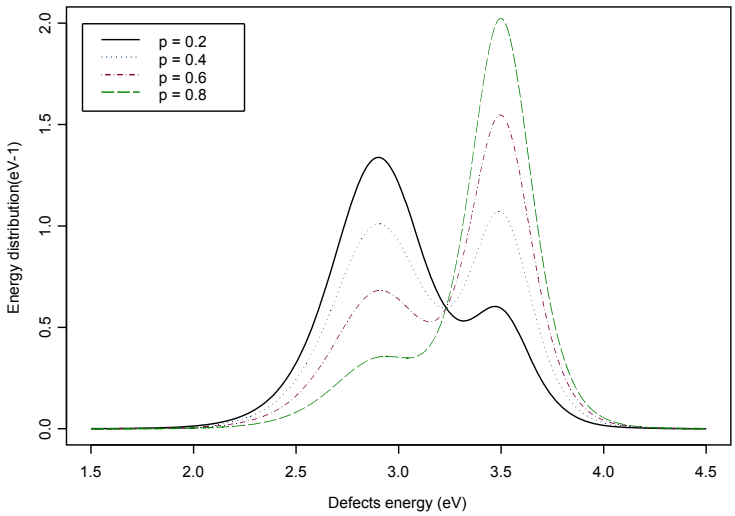


Figure 4: Energy to release hydrogen to different locations (courtesy of Tuttle et al [28]).

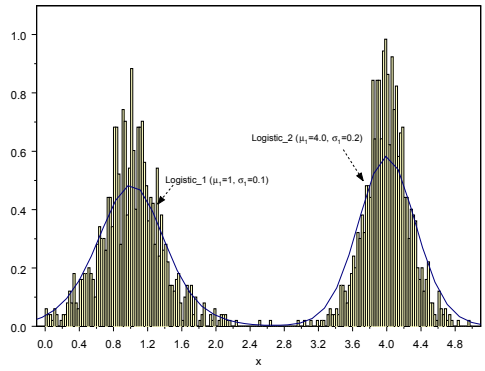


(a)

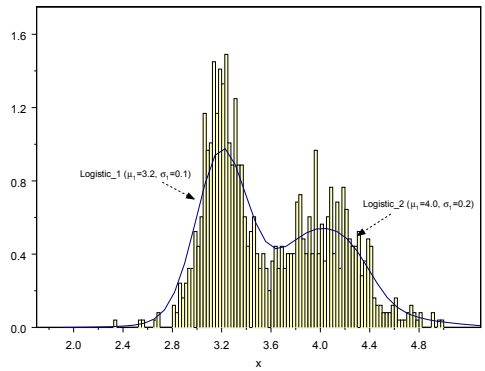


(b)

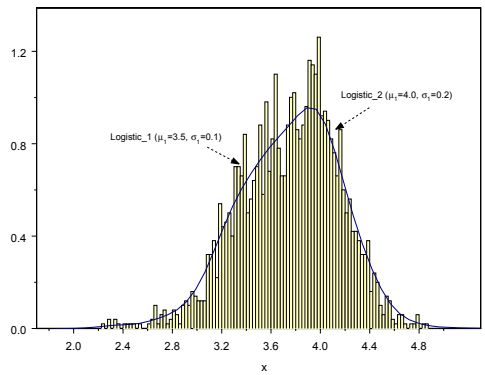
Figure 5: (a)-Defects degradation model (10) and (b)-Mixture distribution of activation energies for defects.



(a) Complete separation



(b) Moderate overlap



(c) Large overlap

Figure 6: Simulation of mixture of two logistic components with varying degrees of separation ($p = 0.5$).

Table 1: Simulation results for mixtures of logistic components: (∗) result tends to a pathological solution.

			Simplex	Direct	Conjugate Gradient	Simulated Annealing
		$\hat{p}$	0.2135 (0.0676)	0.2000 (0.0000)	0.0623 (0.6887)	0.2024 (0.0118)
		$\hat{\mu}_1$	0.9798 (0.0202)	0.9845 (0.0155)	3.4516 (2.4516)	0.9846 (0.0154)
	$p = 0.2$	$\hat{\sigma}_1$	0.0649 (0.3508)	0.0701 (0.2991)	(*)	0.0701 (0.2986)
		$\hat{\mu}_2$	4.0297 (0.0074)	4.0326 (0.0081)	3.2261 (0.1935)	4.0324 (0.0081)
		$\hat{\sigma}_2$	0.2558 (0.2790)	0.2459 (0.2294)	(*)	0.2459 (0.2294)
		$\hat{p}$	0.5380 (0.0759)	0.5000 (0.0000)	0.3000 (0.4000)	0.4988 (0.0025)
		$\hat{\mu}_1$	0.9899 (0.0101)	1.0077 (0.0077)	(*)	1.0077 (0.0077)
Complete Separation:	$p = 0.5$	$\hat{\sigma}_1$	0.1098 (0.0976)	0.1065 (0.0.0652)	(*)	0.1065 (0.0654)
$(\mu_1, \sigma_1, \mu_2, \sigma_2)$ =		$\hat{\mu}_2$	3.9789 (0.0053)	3.9793 (0.0052)	2.9987 (0.2503)	3.9793 (0.0052)
(1.0, 0.1, 4.0, 0.2)		$\hat{\sigma}_2$	0.1686 (0.1570)	0.1878 (0.0609)	(*)	0.1878 (0.0611)
		$\hat{p}$	0.2365 (0.7043)	0.8000 (0.0000)	0.3000 (0.6250)	0.8013 (0.0016)
		$\hat{\mu}_1$	0.9227 (0.0773)	1.0036 (0.0036)	(*)	1.0036 (0.0036)
	$p = 0.8$	$\hat{\sigma}_1$	0.0336 (0.6638)	0.1037 (0.0367)	(*)	0.1036 (0.0364)
		$\hat{\mu}_2$	1.4642 (0.6340)	4.1985 (0.0496)	2.9690 (0.2578)	4.1988 (0.0497)
		$\hat{\sigma}_2$	(*)	0.2399 (0.1997)	(*)	0.2399 (0.1993)
		$\hat{p}$	0.1159 (0.4204)	0.1407 (0.2965)	(*)	0.1401 (0.2994)
		$\hat{\mu}_1$	3.1986 (0.0004)	3.2030 (0.0009)	3.7810 (0.1816)	3.2030 (0.0009)
	$p = 0.2$	$\hat{\sigma}_1$	0.0269 (0.7313)	0.0317 (0.6829)	(*)	0.0318 (0.6820)
		$\hat{\mu}_2$	3.8942 (0.0265)	3.9043 (0.0239)	3.8523 (0.0369)	3.9044 (0.0239)
		$\hat{\sigma}_2$	0.2274 (0.1369)	0.2448 (0.2238)	0.2361 (0.1807)	0.2447(0.2233)
		$\hat{p}$	(*)	0.5354 (0.0707)	0.5994 (0.1989)	0.5358 (0.0717)
		$\hat{\mu}_1$	1.5640 (0.5112)	3.2309 (0.0097)	3.1778 (0.0069)	3.2310 (0.0097)
Moderate overlap:	$p = 0.5$	$\hat{\sigma}_1$	0.0059 (0.9407)	0.1058 (0.0577)	0.0828 (0.1723)	0.1058 (0.0580)
$(\mu_1, \sigma_1, \mu_2, \sigma_2)$=		$\hat{\mu}_2$	3.5660 (0.1086)	4.0194 (0.0049)	4.0274 (0.0068)	4.0194 (0.0049)
(3.2, 0.1, 4.0, 0.2)		$\hat{\sigma}_2$	0.2820 (0.4102)	0.1707 (0.1463)	0.1327 (0.3365)	0.1707 (0.1466)
		$\hat{p}$	0.7890 (0.0138)	0.9900 (0.2375)	0.9132 (0.1415)	0.7380 (0.0776)
		$\hat{\mu}_1$	3.3924 (0.0601)	3.3168 (0.0365)	3.2198 (0.0062)	3.1909 (0.0028)
	$p = 0.8$	$\hat{\sigma}_1$	(*)	0.0993(0.0070)	0.1679 (0.6794)	0.1038 (0.0383)
		$\hat{\mu}_2$	3.1957 (0.2011)	3.8810 (0.0297)	4.4663 (0.1166)	3.9269 (0.0183)
		$\hat{\sigma}_2$	0.0469 (0.7654)	0.2503 (0.2515)	0.3196 (0.5981)	0.2224 (0.1120)
		$\hat{p}$	0.0667 (0.6663)	0.0100 (0.9501)	0.0968 (0.5162)	0.1158 (0.4211)
		$\hat{\mu}_1$	3.3662 (0.0382)	3.7802 (0.0800)	3.7811 (0.0803)	4.6594 (0.3313)
	$p = 0.2$	$\hat{\sigma}_1$	0.0784 (0.2161)	0.2193 (1.1930)	(*)	0.1880 (0.8796)
		$\hat{\mu}_2$	3.9515 (0.0121)	3.9235 (0.0191)	3.8212 (0.0447)	3.8520 (0.0370)
		$\hat{\sigma}_2$	0.2224 (0.1120)	0.2317 (0.1583)	0.3885 (0.9426)	0.1843(0.0784)
		$\hat{p}$	0.6691 (0.3382)	0.5753 (0.1506)	0.6052 (0.2103)	0.5734 (0.1469)
		$\hat{\mu}_1$	3.5146 (0.0042)	3.6536 (0.0439)	3.4518 (0.0138)	3.4909 (0.0026)
Large overlap:	$p = 0.5$	$\hat{\sigma}_1$	0.1309 (0.3085)	0.1124 (0.1240)	0.1227 (0.2269)	0.0892 (0.1078)
$(\mu_1, \sigma_1, \mu_2, \sigma_2)$=		$\hat{\mu}_2$	4.1835 (0.0459)	4.7028 (0.1757)	4.2314 (0.0578)	4.0164 (0.0041)
(3.5, 0.1, 4.0, 0.2)		$\hat{\sigma}_2$	0.1309 (0.3457)	0.1957 (0.0215)	0.3010 (0.5048)	0.2492 (0.2462)
		$\hat{p}$	0.7272 (0.0911)	0.8403 (0.0504)	0.8389 (0.0487)	0.8396 (0.0495)
		$\hat{\mu}_1$	3.6389 (0.0397)	3.5123 (0.0035)	3.5120 (0.0034)	3.5120 (0.0034)
	$p = 0.8$	$\hat{\sigma}_1$	0.2027 (1.0274)	0.1011 (0.0107)	0.1009 (0.0090)	0.1009 (0.0087)
		$\hat{\mu}_2$	3.5063 (0.1234)	4.0418 (0.0105)	4.0389 (0.0097)	4.0378 (0.0095)
		$\hat{\sigma}_2$	0.0405 (0.7973)	0.2912 (0.4558)	0.2907 (0.4534)	0.2904 (0.4521)

Statistical Models for Hot Electron Degradation in Nano-Scaled MOSFET Devices

Suk Joo Bae **Seong-Joon Kim**

Department of Industrial Engineering, Hanyang University, Seoul, Korea

Way Kuo

Department of Electrical & Computer Engineering, The University of Tennessee, Knoxville

Paul H. Kvam

School of Industrial & Systems Engineering, Georgia Institute of Technology

Abstract

In a MOS structure, the generation of hot carrier interface states is a critical feature of the device's reliability. On the nano-scale, there are problems with degradation in transconductance, shift in threshold voltage, and decrease in drain current capability. Quantum mechanics has been used to relate this decrease to degradation and device failure. Although the lifetime and degradation of a device are typically used to characterize its reliability, in this paper we model the distribution of hot-electron activation energies, which has appeal because it exhibits two-point discrete mixture of logistic distributions. The logistic mixture presents computational problems that are addressed in simulation.

Index Terms– EM Algorithm, Logistic Distribution, Maximum Likelihood, Mixture Distribution, Nanotechnology, Reliability.

ACRONYMS

cdf	cumulative distribution function
HCI	hot carrier injection
IC	integrated circuit
MLE	maximum likelihood estimator
$\mathcal{MN}$	multivariate normal
MOS	metal-oxide-semiconductor
MOSFET	metal-oxide-semiconductor field effect transistor
pdf	probability density function

NOTATION

C_i	constant of hot carrier induced degradation model for $i = 1, 2, 3$
$D_{IT}(t_0)$	original interface trap density
$\Delta D_{IT}(t)$	hot carrier activated trap density at time t
$\mathcal{D}(t)$	degradation of a MOSFET device at time t
D_{IT}	interface traps density
E_m	electrical field
$F(\cdot)$	distribution of the hot-electron activation energies
$I(\cdot)$	Fisher information matrix
I_D	drain current
I_{Sub}	substrate current
k	reaction constant
$l(\cdot)$	log-likelihood function
n_b	concentration of Si-H bonds at the interface
n_0	initial concentration of Si-H bonds at the interface
N_T	total concentration of Si bonds
$N_{IT}(t_0)$	initial concentration of interface traps for $t_0 = 0$

$N_{IT}(t)$	concentration of interface traps at time t
$\Delta N_{IT}(t)$	generated interface traps
p	probability of higher activation energies; $p \in (0, 1)$
τ, τ_1, τ_2	lifetime constants
Si^*	Si dangling bond
W	channel width of a device
V_{DD}	power supply voltage
β	coefficient for the I_{Sub}-V_{DD} relationship
μ, σ	parameters of a logistic distribution
μ_i, σ_i	parameters of a mixture logistic distribution for $i = 1, 2$
$\Theta(\Theta_n)$	parameter space (with samples size n)
φ_{IT}	critical energy in electronvolts (eV) for generating an interface trap
φ_0	minimum energy (eV) that an electron must possess to create impact ionization
$\bar{\varphi}_{IT}$, $\bar{\varphi}_{IT,1}$, $\bar{\varphi}_{IT,2}$	mean defect energies
q	elementary charge with the value $1.60218 \times 10^{-19} C$
λ	hot-electron mean-free-path

1 Introduction

The study of reliability has played a vital role in the engineering of products, both large scale and micro scale. In the next decade, it will play an even bigger role for industries in *nanofabrication*, which amounts to designing and manufacturing devices on the nanometer scale; a nanometer (1 nm = one billionth of a meter) is approximately the length of a row of ten hydrogen atoms.

Actually, standard reliability analysis is already essential for the efficient manufacture of nano-devices, but the field of nanotechnology is virtually devoid of results that address reliability issues that are unique to this scale of product. In fact, just as basic physics principles must be rethought at the quantum level, current reliability theories and methods are only partially applicable to systems operating on a nanometer scale. On the molecular level, familiar material properties like conductivity no longer obey laws based on macro scale materials (e.g., Ohm's law). In the same sense, the essential metrics of reliability analysis - material degradation, fatigue, and basic failure mechanisms assume new meaning on the nanometer scale. Sennhauser [25] noted that traditional reliability models may be insufficient due to quantum effects, thermal processes and defect diffusion processes. Experimenters need to consider additional sources of variation such as thermal fluctuations, quantum statistics and Heisenburg uncertainty [3].

There is great potential for reliability improvement if only because current nano-devices are riddled with defects that cause frequent failure problems; the devices are easily damaged by defects that are otherwise harmless to larger micro-devices. A full understanding of the physics and statistics of the defect generation is required in order to investigate the ultimate reliability limitations for nano-devices.

In a MOS structure, for example, the generation of hot carrier interface state is a critical feature of the device's lifetime measurement. Gate current of MOSFETs is made up of electrons injected into the gate oxide by quasi-elastic scattering [13]. However, electrons with high kinetic energies (called "hot carriers") can generate electron-hole pairs near the drain due to impact ionization from atomic-levelled collisions. Those carriers may be injected into the gate oxide and trapped

on defect sites in the oxide. It results in creation of interface states at Si-SiO_2 interface which leads to degradation in transconductance, shift in threshold voltage, and decrease in drain current capability [15]. Understanding the physical mechanisms of HCI will provide meaningful clues for backtracking from observed macro-defects to inferred nano-defects scattered inside the MOSFET devices. This is analogous to reliability problems in which system failure data are used to infer properties about the system's components. In this paper, we investigate physical models of the defects (hot carriers) generation leading to failure based on statistical properties for MOSFETs.

The main results are contained in Section 2. A model for hot electron degradation is achieved via the mixture distribution of hot-electron activation energies. Procedures for statistical inference are outlined in Section 3, and Section 4 contains a discussion of its computation.

2 Physical Models for Hot Carrier Interface State Generation

In a MOS structure, a thin layer of silicon dioxide (SiO_2) forms the insulating layer between the control gate and the conducting channel of transistors used in modern ICs (see Figure 1). As circuits have been made denser to meet the increasing demand for faster logic and memory devices, the dimensions of the transistors have been reduced ("scaled") correspondingly. For example, SiO_2 layer thickness has decreased to 2.0 nm or less, but technology cannot shrink these dimensions indefinitely because thinning down the oxide thickness raises severe technological problems: dielectric thickness variation, penetration of impurities from the highly doped polysilicon gate, reliability and lifetime problems for devices made with the ultrathin oxides, etc [23].

In particular, hot carrier induced degradation in SiO_2 films is perceived as a main potential obstacle for the continued down-scaling of MOSFET devices. During device operation, the film is subjected to electrical stress, and electronic defects like hot carriers that limit device lifetimes are more likely to be created for short-channel devices. Generally, silicon-based transistors are

annealed[1] in a hydrogen-rich environment in order to passivate[2] defects at the Si-SiO_2 interface. However, hydrogen (H) is known to play a key role in the HCI degradation of the transistors with smaller geometries. Lyding et al. [15] proposed to replace hydrogen with deuterium during the final wafer sintering process[3] in order to reduce susceptibility to hot electron degradation effects. The details as to how hydrogen degrades a MOSFET device will be illustrated at the following.

2.1 Mechanisms of Hydrogen Release from the Si-SiO_2 Interface

A principal mechanism of MOSFET degradation is the creation of an interface state (or traps) at the Si-SiO_2 interface. The creation is mainly caused by desorption of hydrogens from the passivated dangling bonds at the mismatched Si-SiO_2 interface. This depassivation is activated directly by the hot electrons that exist during transistor switching. The hot electrons near the drain (see Figure 1) in short-channel devices can generate electron-hole pairs via impact ionization[4] [9]. Figure 2 describes the depassivation procedure at Si-SiO_2 interface. First, electrons (e) or holes (h) with high kinetic energies are attracted to the Si-SiO_2 interface, which weaken the Si-H bond until it breaks. As a result, the hydrogen diffuses into the oxide or Si substrate, subsequently creating interface traps with density D_{IT}. The hydrogen (H: atom and H^+: ion) release reaction breaking Si-H bonds is described in the following equation:

$$Si - H + H^+ + e^- \rightarrow Si^* + H_2 \qquad \text{or} \qquad Si - H + H \rightarrow Si^* + H_2, \tag{1}$$

where Si^* represents the Si dangling bond that is an "interface trap".

The hot-carrier-induced trap density D_{IT} is directly proportional to the concentration of Si dangling bonds at the interface. Because the amount of degradation of a MOSFET switching current that leads to device failure is a function of the variation in the interface trap density activated by the hot carriers, the amount of degradation of a MOSFET device can be represented

[1]The annealing is a process that the transistors are heated at sufficiently high temperatures and slowly cooled down

[2]To treat a subject in order to reduce the chemical reactivity

[3]A process of forming a coherent mass by heating without melting

[4]The formation of or separation into ions by heat, electrical discharge, radiation, or chemical reaction

as the concentration of Si dangling bonds at the interface, which can be measured by D_{IT} via charge pumping technique [8].

Electrons need sufficient activation energies to surmount a surface energy barrier to generate interface traps, and the activation energies are directly linked to interface trap defects. Defects do not necessarily have the same activation energy; in fact, there exists a distribution of activation energies for the HCI generation failure mechanism [17]. From a chemistry viewpoint, there are two reasons to expect a distribution of activation energies for MOSFET devices: the variation in the bond energies due to Si-SiO_2 interface disorder, and the possibility of multiple pathways to activation. If the bonding energies are homogeneously distributed at the interface, the activation energy distribution will be of unimodal form, but the distribution will be bimodal if there exist competing mechanisms of interface state formation following multiple pathways to activation [9].

2.2 Activation Energy Distribution of Hot Carrier Induced Defects

Because the activation of hydrogen at the passivated Si-SiO_2 interface is caused by collisions with electrons (or holes) flowing in the channel, it is crucial to identify the energy distribution of these electrons as a function of the number of interface traps over time to evaluate reliability of MOSFET devices.

N_{IT} is proportional to the concentration of the Si dangling bonds at the interface $N_T - n_b$, where N_T is a total concentration of Si bonds which are able to appear as dangling ones if hydrogen leaves the bond. The time dependent trap generation can be described by a simple version of power law [19]:

$$\Delta N_{IT}(t) = N_{IT}(t) - N_{IT}(t_0) = \frac{n_0}{1 + (kt)^{-\alpha}}. \tag{2}$$

The reaction constant k and the power α are values which can be estimated from the experimental data. In terms of the concentration of interface traps, the degradation of a MOSFET device can be approximated by

$$\mathcal{D}(t) = \frac{\Delta D_{IT}(t)}{D_{IT}(t_0)} \simeq \frac{\Delta N_{IT}(t)}{N_{IT}(t_0)} = \frac{1}{1 + (t/\tau)^{-\alpha}}, \tag{3}$$

where $\tau = 1/k$ is a lifetime constant that has units of time. Note that the model (3) is identical to the degradation model for grating decays in optical interconnects derived from Bragg grating theory in Erdogan et al [6]. For the hot carrier degradation mechanism, lifetime determination is based on the observed accelerated degradation of drain voltage. This is because hot carrier degradation is not accelerated by an increase in temperature [10]. The degradation of a device in terms of $\Delta N_{IT}(t)$ can be related to at I_D [13] as

$$\Delta N_{IT}(t) = C_1 \left[t \frac{I_D}{W} \exp\left(-\frac{\varphi_{IT}}{q\lambda E_m} \right) \right]^n . \tag{4}$$

Introducing an easily measurable I_{Sub} to monitor the device degradation, E_m can be represented with the multiplication factor

$$\frac{I_{Sub}}{I_D} = C_2 \exp\left(-\frac{\varphi_0}{q\lambda E_m} \right). \tag{5}$$

A lifetime is defined as the time to reach a fixed number of interface traps. By combining (4) and (5),

$$\frac{tI_D}{W} \propto \left[\frac{I_{Sub}}{I_D} \right]^{-\varphi_{IT}/\varphi_0} \qquad \text{or} \qquad t \propto I_{Sub}^{-\varphi_{IT}/\varphi_0}. \tag{6}$$

The substrate current is a function of the power supply voltage as

$$I_{Sub} \propto \exp\left(-\frac{\beta}{V_{DD}} \right). \tag{7}$$

Combining (6) and (7), the lifetime of the device can be written as

$$t = C_3 \exp\left(\frac{\varphi_{IT}\beta}{\varphi_0 V_{DD}} \right), \tag{8}$$

and replacing $C_3 \exp(\bar{\varphi}_{IT}\beta/(\varphi_0 V_{DD}))$ by experimentally observed τ in (3), finally the degradation of a MOSFET device can be approximated by using the following distribution on the hot-electron activation energies:

$$F(\varphi_{IT}) = \left[1 + \exp\left(-\frac{\varphi_{IT} - \bar{\varphi}_{IT}}{\sigma} \right) \right]^{-1}, \tag{9}$$

where $\sigma = (\varphi_0 V_{DD})/(\alpha\beta)$. Note that the degradation model $\mathcal{D}(t)$ in (3) that is represented as the proportion of activated defects before time t is equivalent to the probability that activated defects have activation energy less than or equal to φ_{IT}.

Although the lifetime and degradation of a device are typically used to characterize its reliability, in this case the distribution of hot-electron activation energies has a unique appeal because it has a common logistic structure. The logistic distribution, derived from disorder-induced variations in the *Si-H* activation energies, is identical to a Fermi-derivative distribution of the energies of electronic states [5]. Figure 4-(a) shows how the degradation model caused by interface traps (or defects) varies as a function of power supply voltage(V_{DD}) of a MOSFET device for fixed values of β and φ_0, along with the distributions of defect activation energies in 4-(b). The parameter values in the figure are simulated from experimental observations of short-time-tests for 180 nm MOSFET devices in Haggag et al. [9].

2.3 Bimodal Distribution of Activation Energies

Existence of multiple paths and competing mechanisms for the release of hydrogen yields inhomogeneous activation energy distributions. Figure 3 shows the energy level of hydrogen release to different activation pathways (the hydrogen may be attracted to Si or SiO_2 or the Si-SiO_2 interface). Through atomic simulations based on density functional theory, Tuttle et al. [28] showed that the activation energy of hydrogen is distributed around 3.5 eV if the hydrogen desorbs into the SiO_2, but below 3 eV if the final hydrogen state is closer to the silicon bulk. As a result, the time-dependent HCI degradation model is a *mixture* of the model (3) [11]:

$$\mathcal{D}(t) = \frac{\Delta D_{IT}(t)}{D_{IT}(t_0)} = \frac{p}{1 + (t/\tau_1)^{-\alpha_1}} + \frac{1-p}{1 + (t/\tau_2)^{-\alpha_2}}. \tag{10}$$

By letting $\tau_1 = C_{3,1} \exp(\bar{\varphi}_{IT,1}\beta_1/(\varphi_0 V_{DD}))$ and $\tau_2 = C_{3,2} \exp(\bar{\varphi}_{IT,2}\beta_2/(\varphi_0 V_{DD}))$ in (10), the degradation model of a MOSFET device can be represented through φ_{IT} as a mixture of logistic distributions:

$$F(\varphi_{IT}) = p \cdot \left[1 + \exp\left(-\frac{\varphi_{IT} - \bar{\varphi}_{IT,1}}{\sigma_1}\right)\right]^{-1} + (1-p) \cdot \left[1 + \exp\left(-\frac{\varphi_{IT} - \bar{\varphi}_{IT,2}}{\sigma_2}\right)\right]^{-1}, \tag{11}$$

where $\sigma_1 = (\varphi_0 V_{DD})/(\alpha_1\beta_1)$ and $\sigma_2 = (\varphi_0 V_{DD})/(\alpha_2\beta_2)$.

Tuttle et al. [28] experimentally observed a higher mean energy $\bar{\varphi}_{IT,1} \approx 3.5$ eV as well as a lower mean energy $\bar{\varphi}_{IT,2} \approx 2.9$ eV. The higher energy band comes from "single collisions" with

higher energetic electrons and a consequent release of the hydrogen through a higher energy path in the MOSFET. On the other hand, the lower energy band comes from "multiple collisions" with lower energetic electrons and a consequent release of the hydrogen through a lower energy path in the MOSFET [9]. Figures 5-(a) and 5-(b) display the mixture of time-dependent HCI degradation model (10) and the mixture distribution of defect activation energies, respectively, at varying p values with $\bar{\varphi}_{IT,1} = 3.5$ eV, $\bar{\varphi}_{IT,2} = 2.9$ eV, and $V_{DD} = 3.0V$.

3 Parameter Estimation

In this section, we outline the procedures for statistical inference for different characteristics of the MOSFET lifetime. Using the measurement of hot-electron activation energies, we rely on the method of maximum likelihood to estimate logistic model parameters, or more precisely, parameters for the logistic mixture distribution. While the inference for the logistic distribution is straightforward, there are important issues in dealing with estimation for the mixture distribution.

3.1 Logistic Mixture Distribution

The cdf of the random variable X having the logistic distribution is given by

$$F(x;\boldsymbol{\theta}) = \frac{1}{\left\{1+\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}}, \qquad -\infty < x < \infty, \tag{12}$$

for $\boldsymbol{\theta} = (\mu, \sigma)^T$, where μ and σ are location and scale parameters. The corresponding pdf is

$$f(x;\boldsymbol{\theta}) = \frac{1}{\sigma}\frac{\exp\left(-\frac{x-\mu}{\sigma}\right)}{\left\{1+\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}^2} = \frac{1}{\sigma}\left\{F(x)[1-F(x)]\right\}. \tag{13}$$

The pdf of the logistic distribution is symmetric and bell-shaped like that of the normal distribution. Since the logistic distribution has slightly longer tails, it would require an extremely large number of observations to accurately assess whether data come from a normal or logistic distribution. The logistic random variable X has mean $E[X] = \mu$, variance $\mathrm{Var}(X) = (\pi^2\sigma^2)/3$ and coefficient of variation $\pi\sigma/(\mu\sqrt{3})$. The log-likelihood function for a sample of size n from the logistic distribution

is given by

$$l(x_1,\ldots,x_n;\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f(x_i;\boldsymbol{\theta}) = -n\log\sigma - \sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right) - 2\sum_{i=1}^{n}\log\left\{1+\exp\left(-\frac{x_i-\mu}{\sigma}\right)\right\},$$

and the MLEs, $(\hat{\mu},\hat{\sigma})$ of the parameters (μ,σ) satisfy the following likelihood equations:

$$\begin{aligned} \sum_{i=1}^{n}\left\{1+\exp\left(-\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right)\right\}^{-1} &= \frac{n}{2}, \\ -\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right) + \sum_{i=1}^{n}\left(\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right)\left\{1+\exp\left(-\frac{x_i-\hat{\mu}}{\hat{\sigma}}\right)\right\}^{-1} &= \frac{n}{2}. \end{aligned} \tag{14}$$

Taking advantage of the similarity in shape between the logistic and normal distributions, initial values of $\hat{\mu}$ and $\hat{\sigma}$ might be taken as $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$, and $\sqrt{n^{-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}$, respectively. Then solutions could be improved by applying the Newton-Raphson method. When both μ and σ are unknown, the Newton-Raphson method converges quickly to the solutions $\hat{\boldsymbol{\theta}} = (\hat{\mu},\hat{\sigma})^T$. Since the logistic-likelihood function is quasi-concave, the solutions are unique for distinct values of x_i. [1].

The MLEs $\hat{\boldsymbol{\theta}} = (\hat{\mu},\hat{\sigma})^T$, as consistent roots of the likelihood equations (14), satisfy

$$\left(\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)\right) \xrightarrow{L} \mathcal{MN}(\mathbf{0}, I^{-1}(\boldsymbol{\theta}_0)), \tag{15}$$

where $\boldsymbol{\theta}_0$ is the true value of $\boldsymbol{\theta}$ and the Fisher information $I(\boldsymbol{\theta}_0)$ is given by

$$I(\boldsymbol{\theta}_0) = -E\begin{bmatrix} \frac{\partial^2 l}{\partial^2\mu^2} & \frac{\partial^2 l}{\partial\mu\partial\sigma} \\ \frac{\partial^2 l}{\partial\mu\partial\sigma} & \frac{\partial^2 l}{\partial^2\sigma^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{3\sigma^2} & 0 \\ 0 & \frac{3+\pi^2}{9\sigma^2} \end{bmatrix}. \tag{16}$$

It is common in practice to estimate the inverse of the covariance matrix of the MLE by the observed information matrix $I(\hat{\boldsymbol{\theta}})$ rather than the expected information matrix $I(\boldsymbol{\theta}_0)$ evaluated at $\boldsymbol{\theta}_0 = \hat{\boldsymbol{\theta}}$. In general, the observed information matrix is more convenient to use than the expected information matrix as it does not require an expectation to be taken. However as shown in (16), the expectations are trivial in the logistic case and we can easily derive the covariance matrix of the MLE from the expected information matrix. When only μ is unknown, a MLE $\hat{\mu}$ can be uniquely

determined by replacing $\hat{\sigma}$ with known σ value in the likelihood equation. Alternatively, we can use an estimator

$$\zeta_n = \bar{X}_n - \frac{l'(\bar{X}_n)}{l''(\bar{X}_n)}, \tag{17}$$

instead of $\hat{\mu}$ because they have the same asymptotic distribution. Here $\bar{X}_n$ is the average of n samples as a $\sqrt{n}$-consistent estimator (denoted as $\tilde{\mu}$) and

$$l'(\bar{X}_n) = \frac{\partial}{\partial \mu} l(\mathbf{x}; \boldsymbol{\theta})\big|_{\mu=\bar{X}_n} = n - 2\sum_{i=1}^{n} \frac{\exp\left(\frac{x_i-\bar{x}_n}{\sigma}\right)}{1+\exp\left(\frac{x_i-\bar{x}_n}{\sigma}\right)},$$
$$l''(\bar{X}_n) = \frac{\partial^2}{\partial \mu^2} l(\mathbf{x}; \boldsymbol{\theta})\big|_{\mu=\bar{X}_n} = -2\sum_{i=1}^{n} \frac{\exp\left(\frac{x_i-\bar{x}_n}{\sigma}\right)}{1+\exp\left(\frac{x_i-\bar{x}_n}{\sigma}\right)^2}.$$

Theorem 3.1 *Let $\hat{\mu}$ be the MLE of μ, and let ζ_n be given by (17), then $\sqrt{n}(\zeta_n - \hat{\mu}) \to 0$ as $n \to \infty$.*

The proof is listed in the appendix. When only μ is known, the MLE $\hat{\sigma}$ can be uniquely determined by replacing $\hat{\mu}$ with known μ value in the likelihood equations. With moment estimator $\tilde{\sigma} = \{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2/n\}^{1/2}$, which is a $\sqrt{n}$-consistent estimator of σ,

$$\eta_n = \tilde{\sigma} - \frac{l'(\tilde{\sigma})}{l''(\tilde{\sigma})} \tag{18}$$

has the same asymptotic distribution as the MLE $\hat{\sigma}$ (from the theorem) with $\tilde{\sigma}$ instead of $\bar{X}_n$. Here,

$$l'(\tilde{\sigma}) = \frac{\partial}{\partial \sigma} l(\mathbf{x}; \boldsymbol{\theta})\big|_{\sigma=\tilde{\sigma}} \qquad \text{and} \qquad l''(\tilde{\sigma}) = \frac{\partial^2}{\partial \sigma^2} l(\mathbf{x}; \boldsymbol{\theta})\big|_{\sigma=\tilde{\sigma}}.$$

The degradation model of a MOSFET device can represented in (11) is a two-point discrete mixture of logistic distributions. For a random variable X generated from this mixture of logistic distributions, then in terms of $\boldsymbol{\theta}_1 = (\mu_1, \sigma_1)^T$, $\boldsymbol{\theta}_2 = (\mu_2, \sigma_2)^T$ and $\boldsymbol{\Psi} = (p, (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T))^T$, X has pdf

$$f(x; \boldsymbol{\Psi}) = p \cdot f_1(x; \boldsymbol{\theta}_1) + (1-p) \cdot f_2(x; \boldsymbol{\theta}_2), \tag{19}$$

and corresponding CDF

$$F(x; \boldsymbol{\Psi}) = p \cdot F_1(x; \boldsymbol{\theta}_1) + (1-p) \cdot F_2(x; \boldsymbol{\theta}_2), \tag{20}$$

where $F_j(x; \boldsymbol{\theta}_j)$ and $f_j(x; \boldsymbol{\theta}_j)$ are from (12) and (13), respectively, with parameters μ_j and σ_j for $j = 1, 2$.

4 Solving the MLE

Several methods have been proposed to estimate the parameter $\mathbf{\Psi}$. The MLE, based on maximizing the log-likelihood function

$$l(x_1, \ldots, x_n; \mathbf{\Psi}) = \sum_{i=1}^{n} \log \left\{ p \cdot f_1(x_i; \boldsymbol{\theta}_1) + (1-p) \cdot f_2(x_i; \boldsymbol{\theta}_2) \right\},$$

possesses a number of desirable statistical properties. The MLE can be solved via the likelihood equation $\partial l(x_1, \ldots, x_n; \mathbf{\Psi})/\partial \mathbf{\Psi} = \mathbf{0}$. The resulting solution $\hat{\mathbf{\Psi}}$ satisfies

$$\hat{p} = \sum_{i=1}^{n} \kappa(x_i; \hat{\mathbf{\Psi}})/n \qquad \text{and} \qquad \sum_{j=1}^{2} \sum_{i=1}^{n} \kappa(x_i; \hat{\mathbf{\Psi}}) \partial \log f_j(x_i; \hat{\boldsymbol{\theta}}_j)/\partial \boldsymbol{\theta} = \mathbf{0}, \tag{21}$$

where $\kappa(x_i; \hat{\mathbf{\Psi}}) = \hat{p} f_1(x_i; \hat{\boldsymbol{\theta}}_1)/[\hat{p} f_1(x_i; \hat{\boldsymbol{\theta}}_1) + (1-\hat{p}) f_2(x_i; \hat{\boldsymbol{\theta}}_2)]$ is the posterior probability that x_i belongs to the first component of the mixture.

The EM algorithm (see McLachlan and Krishnan [16], for example) can be used to find the MLE for mixtures by solving the likelihood equations (21) iteratively. To apply the EM method, we imagine each observation from the mixture distribution comes with an indicator variable that tells us which of the two logistic distributions the observation was generated. In this case, such an indicator is treated as a missing value.

Starting from an arbitrary initial guess, the algorithm operates in two repeated steps. The E-step estimates missing values as they appear in the log-likelihood, then the M-step finds a local optimum to the likelihood using the estimated data in place of what was missing. With the estimated indicator functions, the MLE is solved more simply using (14) in the M-step (as discussed in the last section). Convergence properties for finite mixture models are discussed in Tanaka and Takemura [22]. In this case, because of the heavy tails in the logistic density and the large variances in (16), the convergence can be slow, depending on how good the initial guess is.

In Table 1, we present summary results of the simulation comparing the performance of different MLE/EM algorithms for mixtures of logistic components. Comparisons between varying degrees of separation in the mixture distributions (complete separation, moderate overlap, and large overlap) are illustrated in Figures 6 (a) - (c), respectively. In these simulations, the mixing proportion p

takes on the values .20, .50, and .80. For a given mixture, the component distributions differ from each other only by location and scale differences. For each set of parameter configurations, samples of size $n = 1,000$ were generated from the corresponding mixture of logistic distributions.

Data generation and parameter estimation for logistic mixture were executed using the *mle* program [12], and for maximizing the likelihood function in the EM algorithm, four different methods were used; simplex, direct, conjugate gradient, and simulated annealing (see [20] for details of the four methods).

Each of the methods has strengths and weakness for different types of functions. In the case of the logistic likelihood, they have peculiar differences; Table 1 shows a summary of the simulation comparing the performance of MLE/EM algorithms for mixtures of logistic components. The computation stopped when the relative errors ϵ, of all 5 parameters, $\mathbf{\Psi} \equiv (p, \mu_1, \sigma_1, \mu_2, \sigma_2)'$ reached $10^{-4} \cdot \epsilon \equiv [\mathbf{\Psi}^{(h+1)} - \mathbf{\Psi}^{(h)}]/\mathbf{\Psi}^{(h)}$. The symbol $(*)$ represents result tends to a pathological solution and the relative bias, $|y - \hat{y}|/y$ is calculated over every parameter in parentheses. The direct and simulated annealing methods provide better results than the other two methods for this parameter estimation. However, simulated annealing requires longer process time for the annealing function. In general, the estimation precision increases when the mixing proportion is neutral ($p = 0.5$).

One way of obtaining standard errors of the estimates of the parameters in a mixture model is to approximate the covariance matrix of $\hat{\mathbf{\Psi}}$ by the inverse of the observed information matrix. For mixture models, the sample size has to be very large to guarantee the asymptotic theory of maximum likelihood, hence a resampling approach such as bootstrapping method can be considered to construct standard errors of the estimates of the parameters. Standard error estimation of $\hat{\mathbf{\Psi}}$ can implemented via a bootstrap procedure; Chapter 13 of [18] for an analogous resampling approach.

5 Discussion

Statistical models based on known physical principles (e.g., the power-rule model, the Arrhenius rule, Eyring Model, etc.) have provided strong methods for parametric inference for various testing

problems in manufacturing. Nano-manufacturing will provide more twists to these traditional models due to the nature of nano-defects and Heisenberg uncertainty. This paper provides basic physical modeling for MOSFET devices based on the nano-level degradation that takes place at defect sites in the MOSFET gate oxide. The distribution of hot-electron activation energies proves to be more accessible than analogous measures of degradation or lifetime, and is derived as a logistic mixture distribution using physical principles on the nanoscale. Although the inference problem is ridden with computational challenges, the derivation of MLEs is straightforward using the EM Algorithm.

6 Appendix

Proof of Theorem : For the true value of μ (denoted as μ_0), Taylor expansion of $l'(\hat{\mu})$ about $l'(\mu_0)$ is

$$l'(\hat{\mu}) = l'(\mu_0) + (\hat{\mu} - \mu_0)l''(\mu_0) + \frac{1}{2}(\hat{\mu} - \mu_0)^2 l'''(\mu^*),$$

where μ^* lies between μ_0 and $\hat{\mu}$. $l'''(\cdot)$ exists for the three-times differentiable logistic likelihood function with respect to μ, and it is bounded. Since the left side is zero for the MLE $\hat{\mu}$, we have

$$\sqrt{n}(\hat{\mu} - \mu_0) = -\frac{\sqrt{n}l'(\mu_0)}{l''(\mu_0)} + R_n,$$

where

$$R_n = -\frac{\sqrt{n}l'(\mu_0)}{l''(\mu_0)} \left[\frac{1}{1 + \frac{n}{l''(\mu_0)} \frac{1}{2n}(\hat{\mu} - \mu_0) l'''(\mu^*)} - 1 \right].$$

Analogously, Taylor expansion of $l'(\tilde{\mu})$ about $l'(\mu_0)$ can be represented as

$$l'(\tilde{\mu}) = l'(\mu_0) + (\tilde{\mu} - \mu_0)l''(\mu_0) + \frac{1}{2}(\tilde{\mu} - \mu_0)^2 l'''(\mu^*).$$

It follows from (15) that

$$\zeta_n = \tilde{\mu} - \frac{1}{l''(\tilde{\mu})} \left[l'(\mu_0) + (\tilde{\mu} - \mu_0)l''(\mu_0) + \frac{1}{2}(\tilde{\mu} - \mu_0)^2 l'''(\mu^*) \right],$$

and by using the expansion $l''(\tilde{\mu}) = l''(\mu_0) + (\tilde{\mu} - \mu_0)l'''(\mu^*)$,

$$\sqrt{n}(\zeta_n - \mu_0) = -\frac{\sqrt{n}l'(\mu_0)}{l''(\mu_0)} \left(1 + \frac{(\tilde{\mu} - \mu_0)l'''(\mu^*)}{l''(\mu_0)} \right)^{-1} + R'_n,$$

where

$$R'_n = \sqrt{n}(\hat{\mu} - \mu_0)\left[1 - \frac{l''(\mu_0)}{l''(\tilde{\mu})} - \frac{1}{2}(\tilde{\mu} - \mu_0)\frac{l'''(\mu^*)}{l''(\tilde{\mu})}\right].$$

As $n \to \infty$, R_n and R'_n tend to 0 in probability since $\hat{\mu} \to \mu_0$, $\frac{l''(\mu_0)}{l''(\tilde{\mu})} \to 1$, and $\tilde{\mu} \to \mu_0$, and additionally $\left(1 + \frac{(\tilde{\mu}-\mu_0)l'''(\mu^*)}{l''(\mu_0)}\right) \to 1$.

References

[1] Antle, C. E., Klimko, L., and Harkness, W. (1970), "Confidence Intervals for the Parameters of the Logistic Distribution", *Biometrika*, Vol. 57, No. 2, 397–402.

[2] Balakrishnan, N. (1992), *Handbook of the Logistic Distribution*, Dekker, New York.

[3] Birge, R. R., Lawrence, A. F., and Tallent, J. R. (1991), "Quantum effects, thermal statistics and reliability of nanoscale molecular and semiconductor devices", *Nanotechnology*, Vol. 2, 73–87.

[4] DeCani, J. S., and Stine, R. A. (1986), "A Note on Deriving the Information Matrix for a Logistic Distribution", *The American Statistician*, Vol. 40, No. 3, 220–222.

[5] Devine, R. A. B., Autran, J. L., Warren, W. L., VAnheusdan, K. L., and Rostaing, J. C. (1997), "Interfacial Hardness Enhancement in Deuterium Annealed 0.25 μm Channel Metal Oxide Semiconductor Transistors", *Applied Physics Letters*, Vol. 70, No. 22, 2999–3001.

[6] Erdogan, T., Mizrahi, V., Lemaire, P. J., and Monroe, D. (1994), "Decay of Ultraviolet-Induced Fiber Bragg Gratings", *Journal of Applied Physics*, Vol. 76, No. 1, 73–80.

[7] Feng, Z. D., and McCulloch, C. E. (1996), "Using Bootstrap Likelihood Ratios in Finite Mixture Models", *Journal of the Royal Statistical Society B*, Vol. 58, No. 3, 609–617.

[8] Groeseneken, G., Maes, H. E., Beltran, N., and DE Keersmaecker, R. F. (1984), "A Reliable Approach to Charge- Pumping Measurements in MOS Transistors", *IEEE Transactions on Electron Devices*, Vol. 31, No. 1, 42–53.

[9] Haggag, A., McMahon, W., Hess, K., Cheng, K., Lee, J., and Lyding, J. W., (2001), "High-Performance Chip Reliability from Short-Time-Tests: Statistical Models for Optical Interconnect and Deep-Submicron Transistor Failures",*International Reliability Physics Symposium*, Orlando, FL.

[10] Heremans, P., Van den Bosch, G., Bellens, R., Groeseneken, G., Maes, H. E. (1990), "Temperature Dependence of the Channel Hot-Carrier Degradtion of n-Channel MOSFETs", *IEEE Transactions on Electron Devices.*, Vol. 37, No. 4., 980–993.

[11] Hess, K., Haggag, A., McMahon, W., Cheng, K., Cheng, K., Lee, J., and Lyding, J. W. (2001), "The Physics of Determining Chip Reliability", *IEEE Circuits Devices Magazine*, Vol. 17, No. 3, 33–38.

[12] Holman, D. J. (2003), *A Programming Language for Buliding Likelihood Models*, User's manual, Vol. 1, Ver. 2.1.

[13] Hu, C., Tam, S. C., Hsu, F., Ko, P., Chan, T., and Terrill, K. W. (1985), "Hot-Electron-Induced MOSFET Degradation-Model, Monitor, and Improvement', *IEEE Journal of Solid-State Circuits*, Vol. 20, No. 1, 295–305.

[14] Lehmann, E. L., and Casella, G. (1998), *Theory of Point Estimation*, Springer, New York.

[15] Lyding, J. W., Hess, K., and Kizilyalli, I. C. (1996), "Reduction of Hot Electron Degradation in Metal Oxide Semiconductor Transistors by Deuterium Processing", *Applied Physics Letters*, Vol 68. No. 18, 2526–2528.

[16] McLachlan, G. J., and Krishnan, T. (1997), *The EM Algorithm and Extensions*, Wiley, New York.

[17] McMahon, W., Haggag, A., and Hess, K. (2003), "Reliability Scaling Issues for Nanoscale Devices", *IEEE Transactions on Nanotechnology*, Vol. 2, No. 1, 33–38.

[18] Meeker, W. Q. and Escobar, L. A. (1998), *Statistical Methods for Reliability Data*, Wiley, New York.

[19] Penzin, O., Haggag, A., McMahon, W., Lyumkis, E., and Hess, K. (2003), "MOSFET Degradation Kinetics and Its Simulation", *IEEE Transactions on Electron Devices*, Vol. 50, No. 6, 1445–1450.

[20] Press, W. H., Flannery B. P., Teukolsky S. A., and Vetterling W. T. (1989), *Numerical Recipes in Pascal: The Art of Scientific Programming*, Cambridge University Press, Cambridge.

[21] Quandt, R. E., and Ramsey, J. B. (1978), "Estimating Mixtures of Normal Distributions and Switching Regressions", *Journal of American Statistical Association*, Vol. 73, 730-738.

[22] Tanaka, K., and Takemura, A. (2003), "Strong Consistency of MLE for Finite Mixtures of Location-Scale Distributions When the Scale Parameters are Exponentially Small", *Mathematical Engineering Technical Report*, University of Tokyo.

[23] Waser, R., (2003), *Nanoelectronics and Information Technology: Advanced Electronic Materials and Novel Devices.*, Wiley, Weinheim.

[24] Sennhauser, U., Reiner, J., and Nellen, P. M. (2004), "Nanoreliability", *13th NID Workshop*, Athens, Greece.

[25] Sennhauser, U. (2004), "Reliability of Nanostructed Materials and Devices", Project No. 5202.1, Swiss Federal Laboratories for Materials Testing and Research.

[26] Shukla, S. K., Norman, G., Parker, D., and Kwiatkowska, M. (2005), "Evaluating the Reliability of Defect-Tolerant Architectures for Nanotechnology with Probabilistic Model Checking", Submitted for Publication in *IEEE Transactions on Computer Aided Design* (under 2nd revision).

[27] Stathis, J. H. and DiMaria, D. J. (1998), "Reliability Projection for Ultra-Thin Oxides at Low Voltage", *IEDS Technical Digest*, Vol. 6, 167–170.

[28] Tuttle, B., and Van d Walle, C. G. (1999), "Structure, Energetics, and Vibrational Properties of Si-H Bond Dissociation in Silicon", *Physical Review B-Condensed Matter and Materials Physics*, Vol. 39, 12884–12889.

[29] Vallett, D. P. (2002), "Failure Analysis Requirements for Nanoelectronics", *IEEE Transactions on Nanotechnology*, Vol. 1, No. 3, 117–121.

Authors

Suk Joo Bae is an Assistant Professor in the Department of Industrial Engineering at Hanyang University, Seoul, Korea. He received his Ph.D. from the School of Industrial and Systems Engineering at the Georgia Institute of Technology in 2003. He worked as a reliability engineer at the SDI, Korea, from 1996 to 1999. His research interests are centered on reliability evaluation of light displays and nano-devices via accelerated life and degradation testing, statistical robust parameter design, and process control for large-volume on-line processing data. He is a member of INFORMS, ASA, and IMS.

Seong-Joon Kim is currently a Graduate Student in the Department of Industrial Engineering at Hanyang University, Seoul, Korea. His research interests are centered on reliability evaluation of nano-devices, spatial modeling of point process.

Way Kuo is the University Distinguished Professor and Dean of Engineering at the University of Tennessee. He was the Wisenbaker Chair of Engineering and head of the industrial engineering department at Texas A&M. He is an elected member of the National Academy of Engineering. He received his B.S. in nuclear engineering from National Tsing Hua University in Taiwan and his Ph.D in industrial engineering from Kansas State University. He is a fellow of IIE, IEEE, and ASQ, and an Academician of International Academy for Quality. He has chaired the Council of Industrial Engineering Academic Department Heads and the IIE Council of Fellows.

Paul H. Kvam is a Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He received his Ph.D. from the University of California, Davis in 1991. His research interests include reliability evaluation in engineering applications, including accelerated life testing, degradation testing, analysis of dependent systems, and nonparametric inference. He is a fellow of the ASA.

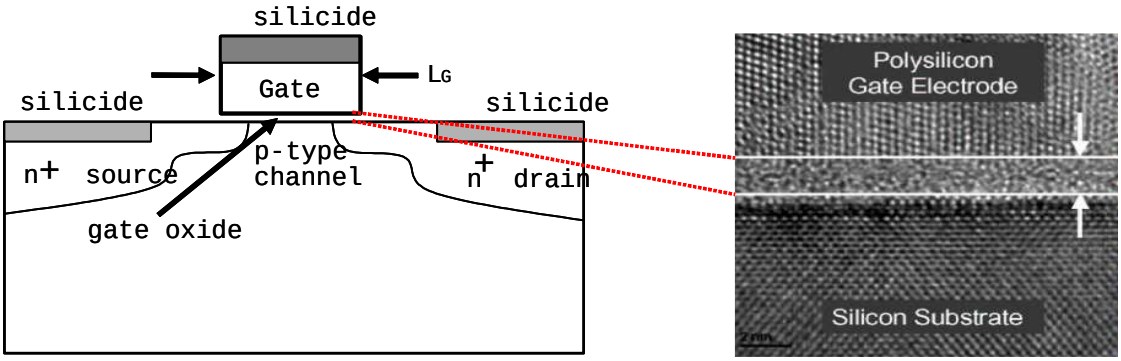


Figure 1: Basic structure of a planar MOSFET: For 90 nm generation gate oxide, the thickness of silicon oxide (SiO_2) is less than 2 nm.

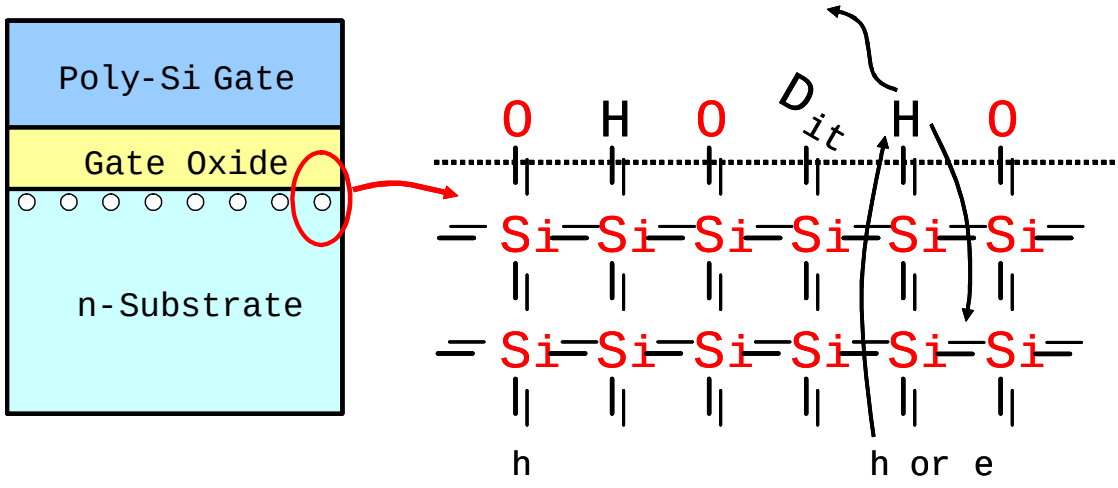


Figure 2: Desorption procedure of hydrogens at Si-SiO_2 interface.

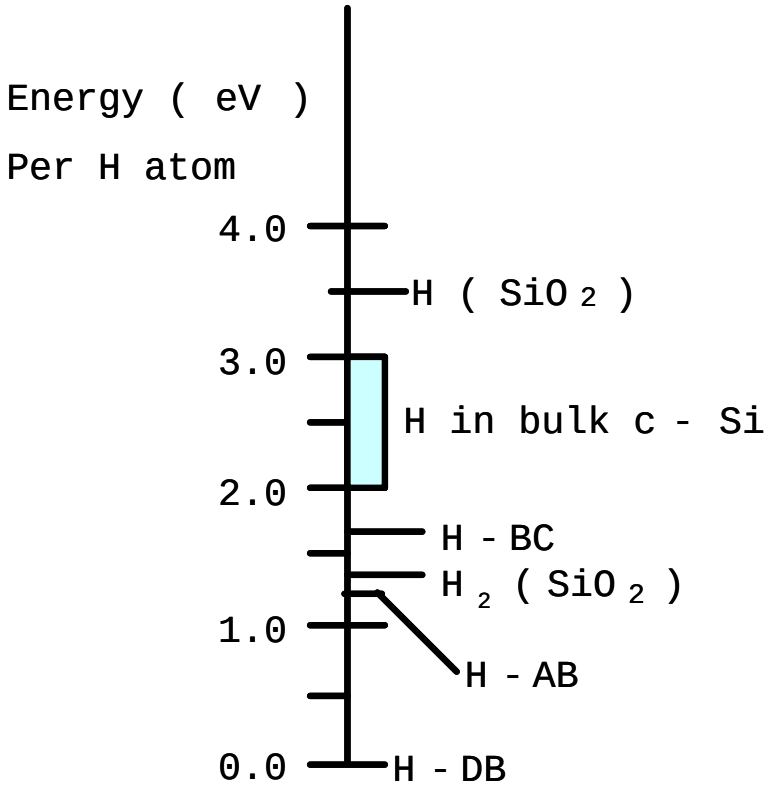
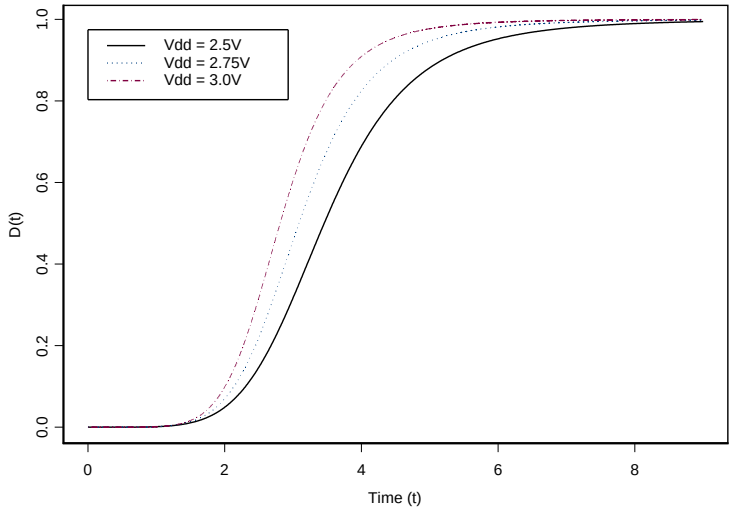
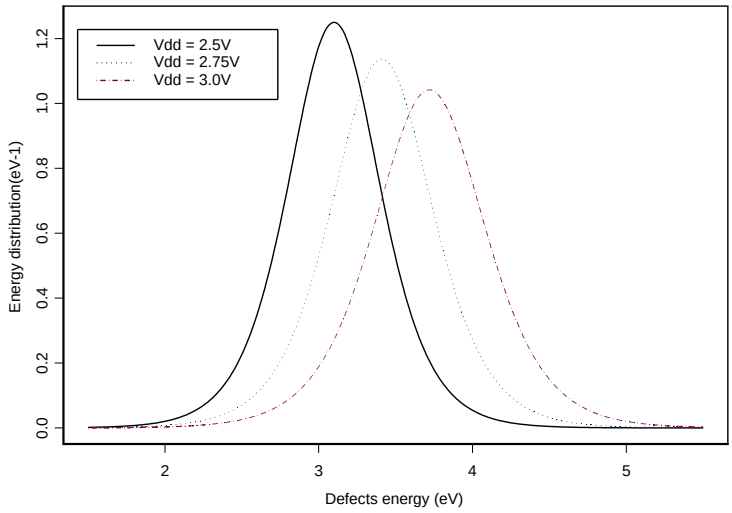


Figure 3: Energy to release hydrogen to different locations (courtesy of Tuttle et al [28]).

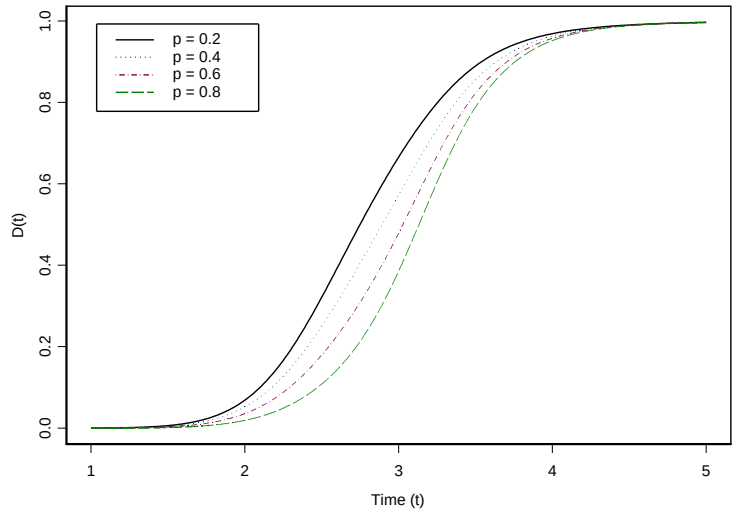


(a)

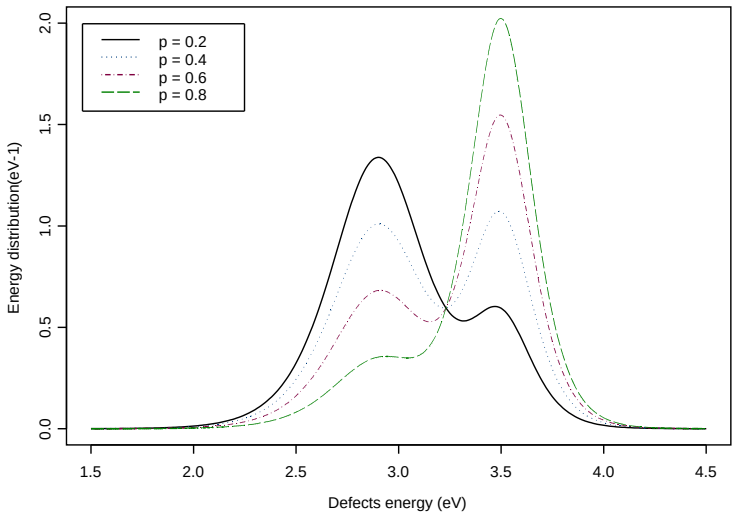


(b)

Figure 4: (a)-Defects degradation model (3); (b)-Distribution of activation energies for interface traps.

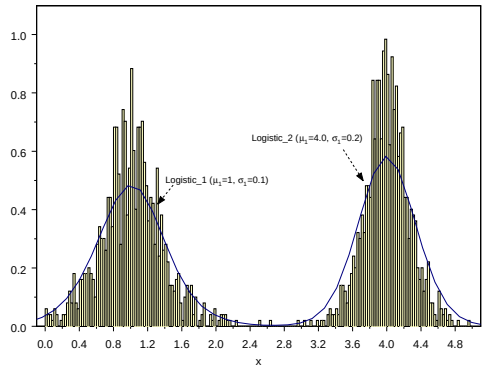


(a)

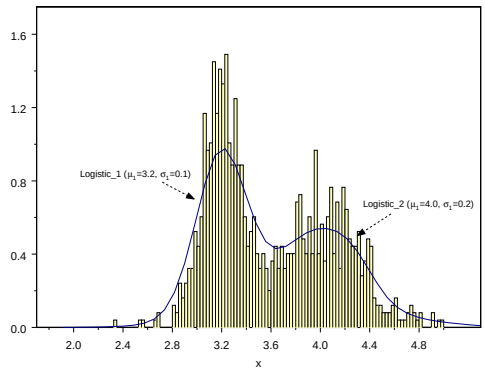


(b)

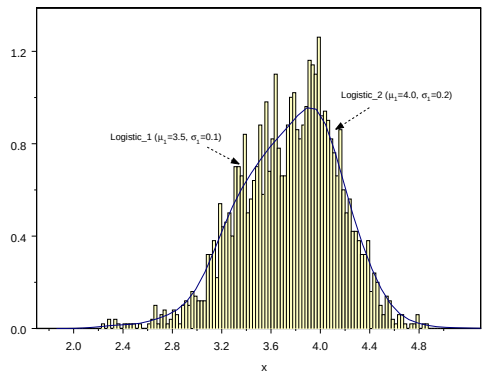
Figure 5: (a)-Defects degradation model (10) and (b)-Mixture distribution of activation energies for defects.



(a) Complete separation



(b) Moderate overlap



(c) Large overlap

Figure 6: Simulation of mixture of two logistic components with varying degrees of separation ($p = 0.5$).

Table 1: Simulation results for mixtures of logistic components: (∗) result tends to a pathological solution.

			Simplex	Direct	Conjugate Gradient	Simulated Annealing
		$\hat{p}$	0.2135 (0.0676)	0.2000 (0.0000)	0.0623 (0.6887)	0.2024 (0.0118)
		$\hat{\mu}_1$	0.9798 (0.0202)	0.9845 (0.0155)	3.4516 (2.4516)	0.9846 (0.0154)
	$p = 0.2$	$\hat{\sigma}_1$	0.0649 (0.3508)	0.0701 (0.2991)	(*)	0.0701 (0.2986)
		$\hat{\mu}_2$	4.0297 (0.0074)	4.0326 (0.0081)	3.2261 (0.1935)	4.0324 (0.0081)
		$\hat{\sigma}_2$	0.2558 (0.2790)	0.2459 (0.2294)	(*)	0.2459 (0.2294)
		$\hat{p}$	0.5380 (0.0759)	0.5000 (0.0000)	0.3000 (0.4000)	0.4988 (0.0025)
		$\hat{\mu}_1$	0.9899 (0.0101)	1.0077 (0.0077)	(*)	1.0077 (0.0077)
Complete Separation:	$p = 0.5$	$\hat{\sigma}_1$	0.1098 (0.0976)	0.1065 (0.0.0652)	(*)	0.1065 (0.0654)
$(\mu_1, \sigma_1, \mu_2, \sigma_2) =$		$\hat{\mu}_2$	3.9789 (0.0053)	3.9793 (0.0052)	2.9987 (0.2503)	3.9793 (0.0052)
(1.0, 0.1, 4.0, 0.2)		$\hat{\sigma}_2$	0.1686 (0.1570)	0.1878 (0.0609)	(*)	0.1878 (0.0611)
		$\hat{p}$	0.2365 (0.7043)	0.8000 (0.0000)	0.3000 (0.6250)	0.8013 (0.0016)
		$\hat{\mu}_1$	0.9227 (0.0773)	1.0036 (0.0036)	(*)	1.0036 (0.0036)
	$p = 0.8$	$\hat{\sigma}_1$	0.0336 (0.6638)	0.1037 (0.0367)	(*)	0.1036 (0.0364)
		$\hat{\mu}_2$	1.4642 (0.6340)	4.1985 (0.0496)	2.9690 (0.2578)	4.1988 (0.0497)
		$\hat{\sigma}_2$	(*)	0.2399 (0.1997)	(*)	0.2399 (0.1993)
		$\hat{p}$	0.1159 (0.4204)	0.1407 (0.2965)	(*)	0.1401 (0.2994)
		$\hat{\mu}_1$	3.1986 (0.0004)	3.2030 (0.0009)	3.7810 (0.1816)	3.2030 (0.0009)
	$p = 0.2$	$\hat{\sigma}_1$	0.0269 (0.7313)	0.0317 (0.6829)	(*)	0.0318 (0.6820)
		$\hat{\mu}_2$	3.8942 (0.0265)	3.9043 (0.0239)	3.8523 (0.0369)	3.9044 (0.0239)
		$\hat{\sigma}_2$	0.2274 (0.1369)	0.2448 (0.2238)	0.2361 (0.1807)	0.2447(0.2233)
		$\hat{p}$	(*)	0.5354 (0.0707)	0.5994 (0.1989)	0.5358 (0.0717)
		$\hat{\mu}_1$	1.5640 (0.5112)	3.2309 (0.0097)	3.1778 (0.0069)	3.2310 (0.0097)
Moderate overlap:	$p = 0.5$	$\hat{\sigma}_1$	0.0059 (0.9407)	0.1058 (0.0577)	0.0828 (0.1723)	0.1058 (0.0580)
$(\mu_1, \sigma_1, \mu_2, \sigma_2)=$		$\hat{\mu}_2$	3.5660 (0.1086)	4.0194 (0.0049)	4.0274 (0.0068)	4.0194 (0.0049)
(3.2, 0.1, 4.0, 0.2)		$\hat{\sigma}_2$	0.2820 (0.4102)	0.1707 (0.1463)	0.1327 (0.3365)	0.1707 (0.1466)
		$\hat{p}$	0.7890 (0.0138)	0.9900 (0.2375)	0.9132 (0.1415)	0.7380 (0.0776)
		$\hat{\mu}_1$	3.3924 (0.0601)	3.3168 (0.0365)	3.2198 (0.0062)	3.1909 (0.0028)
	$p = 0.8$	$\hat{\sigma}_1$	(*)	0.0993(0.0070)	0.1679 (0.6794)	0.1038 (0.0383)
		$\hat{\mu}_2$	3.1957 (0.2011)	3.8810 (0.0297)	4.4663 (0.1166)	3.9269 (0.0183)
		$\hat{\sigma}_2$	0.0469 (0.7654)	0.2503 (0.2515)	0.3196 (0.5981)	0.2224 (0.1120)
		$\hat{p}$	0.0667 (0.6663)	0.0100 (0.9501)	0.0968 (0.5162)	0.1158 (0.4211)
		$\hat{\mu}_1$	3.3662 (0.0382)	3.7802 (0.0800)	3.7811 (0.0803)	4.6594 (0.3313)
	$p = 0.2$	$\hat{\sigma}_1$	0.0784 (0.2161)	0.2193 (1.1930)	(*)	0.1880 (0.8796)
		$\hat{\mu}_2$	3.9515 (0.0121)	3.9235 (0.0191)	3.8212 (0.0447)	3.8520 (0.0370)
		$\hat{\sigma}_2$	0.2224 (0.1120)	0.2317 (0.1583)	0.3885 (0.9426)	0.1843(0.0784)
		$\hat{p}$	0.6691 (0.3382)	0.5753 (0.1506)	0.6052 (0.2103)	0.5734 (0.1469)
		$\hat{\mu}_1$	3.5146 (0.0042)	3.6536 (0.0439)	3.4518 (0.0138)	3.4909 (0.0026)
Large overlap:	$p = 0.5$	$\hat{\sigma}_1$	0.1309 (0.3085)	0.1124 (0.1240)	0.1227 (0.2269)	0.0892 (0.1078)
$(\mu_1, \sigma_1, \mu_2, \sigma_2)=$		$\hat{\mu}_2$	4.1835 (0.0459)	4.7028 (0.1757)	4.2314 (0.0578)	4.0164 (0.0041)
(3.5, 0.1, 4.0, 0.2)		$\hat{\sigma}_2$	0.1309 (0.3457)	0.1957 (0.0215)	0.3010 (0.5048)	0.2492 (0.2462)
		$\hat{p}$	0.7272 (0.0911)	0.8403 (0.0504)	0.8389 (0.0487)	0.8396 (0.0495)
		$\hat{\mu}_1$	3.6389 (0.0397)	3.5123 (0.0035)	3.5120 (0.0034)	3.5120 (0.0034)
	$p = 0.8$	$\hat{\sigma}_1$	0.2027 (1.0274)	0.1011 (0.0107)	0.1009 (0.0090)	0.1009 (0.0087)
		$\hat{\mu}_2$	3.5063 (0.1234)	4.0418 (0.0105)	4.0389 (0.0097)	4.0378 (0.0095)
		$\hat{\sigma}_2$	0.0405 (0.7973)	0.2912 (0.4558)	0.2907 (0.4534)	0.2904 (0.4521)

Length Bias in the Measurements of Carbon Nanotubes

Paul KVAM

H. Milton Stewart School of Industrial Engineering
Georgia Institute of Technology
Atlanta, GA

To measure carbon nanotube lengths, atomic force microscopy and special software are used to identify and measure nanotubes on a square grid. Current practice does not include nanotubes that cross the grid, and, as a result, the sample is length-biased. The selection bias model can be demonstrated through Buffon's needle problem, extended to general curves that more realistically represent the shape of nanotubes observed on a grid. In this article, the nonparametric maximum likelihood estimator is constructed for the length distribution of the nanotubes, and the consequences of the length bias are examined. Probability plots reveal that the corrected length distribution estimate provides a better fit to the Weibull distribution than the original selection-biased observations, thus reinforcing a previous claim about the underlying distribution of synthesized nanotube lengths.

KEY WORDS: Buffon's needle problem; Censoring; Goodness of fit; Nonparametric; Probability plot; Sampling; Selection bias; Weibull; Weighted distributions.

1. INTRODUCTION

Carbon nanotubes are rolled-up nanoscale sheets of graphitic carbon used to enhance materials with tensile strength, thermal conductivity, and electronic conductivity. In terms of tensile strength alone, they can be stiffer and stronger than any other microscaled fiber. Nanotubes and similar nanosized devices have garnered great attention in the industrial and academic world, even though statistical research in nanomanufacturing is still in its early stages. In contrast to laboratory research, commercial applications have developed at a slow pace, due primarily to the enormous production costs of high-quality nanotubes.

A major challenge in the production of nanotubes is controlling and analyzing the important qualities, such as length, diameter, and strength. Nanotube length can be measured using atomic force microscopy (AFM) and special software developed by SIMAGIS, which processes three-dimensional images of bulk nanotube material samples. The software identifies and classifies nanotubes and then analyzes nanotube lengths while recognizing the intersections, contacts, and three dimensions of curvature of the tubes. The solution features also include the ability to filter objects according to selected length and height. In this article we focus on length measurements. Figure 1 presents an example of an AFM profile image of carbon nanotubes in a 5 μm × 5 μm square interval. The average size of these nanotubes is nearly .5 μm, or 500 nm. This recent technological innovation gives experimenters the ability to ascertain distributional properties of carbon nanotubes beyond the simple mean lengths, as characterized by Ziegler et al. (2005). Important mechanical properties that are attainable include tensile strength, elastic modulus, and fracture toughness.

Wang, Liang, Wang, and Zhang (2006) presented a statistical characterization of nanotube length using an AFM sample and the SIMAGIS software. The set of length measurements for 651 detected nanotubes is summarized in a histogram in Figure 2. Like the image in Figure 1, the sampling area is a square grid with sides measuring 5 μm. Although it has been verified in smaller samples that the software accurately finds and measures nanotubes smaller than 10 nanometers, the software's count does not include nanotubes that cross the boundary of the square, and thus the data are length-censored. Although the censoring mechanism can be easily included in the statistical analysis, this sampling method introduces size bias into the analysis as well.

2. LENGTH BIAS

For a positive random variable X with density function $f(x)$, if the random variable is observed with a probability proportional to its size $w(x)$, then the distribution of the random sample obtained with this bias has density

$$f_w(x) = \frac{w(x)f(x)}{\int w(x)f(x)\,dx}.$$

This kind of weighted bias in sampling was first discussed by Rao (1965), with a particular interest in length bias, where $w(x) = x$. This holds for discrete distributions as well, and we use $dF(x)$ to denote the more general class of distributions.

Length bias is a common phenomenon in survival studies in which prevalent cases are determined through cross-sectional study; for example, a test tends to detect more slow-growing cancers that take longer to become symptomatic compared with aggressive, fast-growing cancers. Cancers that grow slowly are easier to detect because they have a longer presymptomatic period when they are detectable, and thus show up with more relative frequency in such cancer studies. Horvath (1985) and Asgharian (2002) have presented a more detailed discussion on sampling and analysis problems induced by length bias.

TECHNOMETRICS, NOVEMBER 2008, VOL. 50, NO. 4
DOI 10.1198/004017008000000442

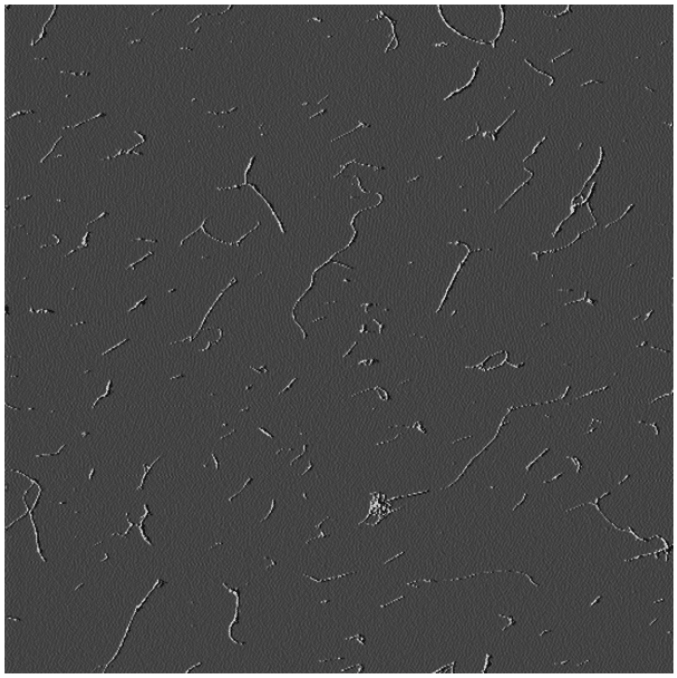


Figure 1. AFM image of carbon nanotubes.

2.1 Buffon's Needle Problem

Buffon's famous needle problem can be used to illustrate the length bias that occurs in this kind of two-dimensional sample. In 1733, Georges-Louis Buffon, a French naturalist and mathematician, posed the following problem: Given a needle of length r and a grid of parallel lines spaced by an equal distance a, what is the probability that a randomly tossed needle will cross one of these lines? The problem is solved with rudimentary geometry based on an assumption that the center of the needle and its angle with respect to the grid are uniformly distributed. This example has been used as a classic probability problem and as an experimental way of obtaining an estimate of the constant π (see, e.g., Perlman and Wichura 1975). In this article we focus on the case in which $r \leq a$, where the censoring probability is $2r/(a\pi)$.

By 1812, Pierre-Simon Laplace considered a second grid of parallel lines constructed at right angles to the first set of parallel lines and with the same distance a, and found that the crossing probability doubled; that is, the probability of a line crossing on a square boundary is $4r/(a\pi)$. A geometrical proof was offered by Mantel (1952).

Gnedenko (1962) provided a more general extension of the crossing probability to multisided convex polygons under the constraint that the diameter is $<a$. Ramely (1969) reframed this extension in terms of arbitrary curves; this is sometimes referred to as "Buffon's noodle problem." With polygons and crooked lines, the crossing frequency changes, but the number of expected crossings remains constant. To illustrate this, think of a needle bent in half. The probability the folded-over needle crosses a grid line is halved, but the expected number of crossings stays the same, because the the needle will necessarily cross the the grid line twice.

To represent a nanotube of length r, we could treat it as a straight needle with crossing probability of $4r/(a\pi)$ or as a crooked, curving tube that connects to itself, in which case the crossing probability is reduced to $2r/(a\pi)$. As shown in Figure 2, the nanotubes exhibit only a modest curvature, much more like a straight needle than a connected polygon.

To illustrate the effect of curvature in a simple way, we assume that any nanotube consists of two segments of equal length. Figure 3 shows the nanotubes with segment angles of $\theta = \{0, \pi/8, \pi/4\}$. If a nanotube consists of two segments of length $r/2$, then the sum of the three side lengths of the inscribed triangle is $r(1 + \cos(\theta))$. Accordingly, the crossing probability is reduced to

$$\frac{4r}{a\pi}\lambda, \tag{1}$$

where $\lambda = (\cos(\theta) + 1)/2$ represents the shrinkage in the crossing probability due to the bending of the nanotube. For $\theta = \{0, \pi/8, \pi/4\}$, we have $\lambda = \{1, .9619, .8536\}$. It might be sensible to treat λ as a random effect, reflecting the apparent heterogeneity of shapes observed in the sample of synthesized

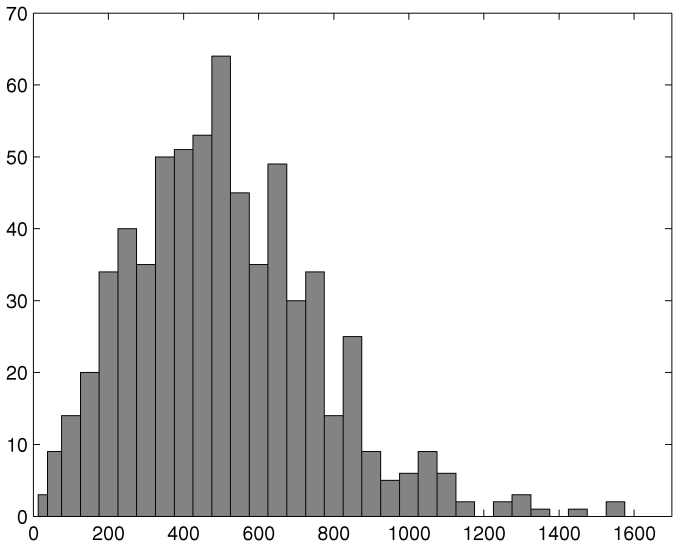


Figure 2. Histogram of dispersed SWNT lengths.

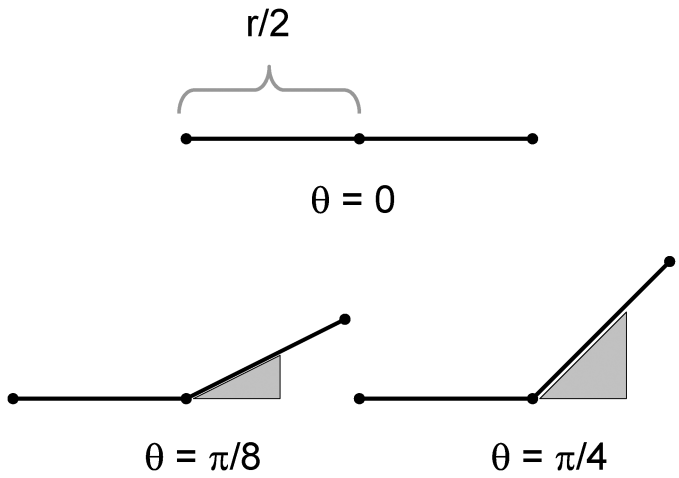


Figure 3. Nanotubes represented as bisected needles with angles $\theta = \{0, \pi/8, \pi/4\}$.

nanotubes. In this article we focus on a fixed, known λ, emphasizing the case of $\lambda = 1$ (straight nanotubes). The effects of curvature and randomness in the shape properties are discussed further in Section 4.

2.2 Censored Nanotubes

Obviously, a longer nanotube is more likely to be censored (and thus not measured) than a smaller one. Consequently, a sample that ignores the censored lines is biased according to the nanotube's length. Suppose that the carbon nanotube lengths have cumulative distribution function $F_L(t)$. We assume $F_L(t_0) = 1$ for some $t_0 < a$, although this constraint is not necessary for demonstrating the selection bias. Let $\mu_L = \int t\, dF_L(t)$. Based on the approximation in (1), the overall probability of censoring is

$$p = \int \frac{4\lambda t}{a\pi}\, dF_L(t) = \frac{4\lambda\mu_L}{a\pi}. \tag{2}$$

From the software output, we observe a sample of n nanotube measurements that has empirical distribution function $\tilde{F}_n(t)$. Let F_0 and F_1 be the distribution functions of the censored observations (nanotubes that cross the boundary) and uncensored observations.

From the mixture

$$F_L(t) = pF_0(t) + (1-p)F_1(t), \tag{3}$$

$\tilde{F}_n(t)$ is actually the nonparametric maximum likelihood estimator (NPMLE) for F_1, not F_L. If the censoring probability p remains small, then the bias from estimating $F_L(t)$ with $\tilde{F}_n(t)$ might be minuscule. For problems in which the nanotube length represents a more significant proportion of a, however, the bias can be sizeable, as we show later. From (3), the length bias $[(w(x) = x)]$ affects both the censored population (F_0) and the observed population (F_1) as

$$\begin{aligned} dF_0(t) &= \frac{t}{\mu_L}\, dF_L(t) \qquad \text{and} \\ dF_1(t) &= \frac{a\pi - 4\lambda t}{a\pi - 4\lambda\mu_L}\, dF_L(t), \end{aligned} \tag{4}$$

where $0 \le t \le a$. Whereas the censored observations are length-biased, the observations lying strictly within the square also are biased, but in an opposite manner.

In the remainder of the article we restrict our consideration to distributions for which both F_0 and F_1 are proper distributions functions. From (4), we require that $F_L(t_0) = 1$ for some $t_0 < a\pi/4$.

2.3 Example

A simple example illustrates how length bias can adversely affect the sample outcome if one does not compensate for it. Suppose that the underlying distribution for (straight) nanotube lengths is Uniform$(0, 2\mu)$, observed on a unit grid $(a = 1)$. From $dF_1(t)$ in (4), the density function of the uncensored nanotubes is

$$f_1(x) = \left(\frac{1}{2\mu}\right)\frac{\pi - 4x}{\pi - 4\mu}, \qquad 0 \le x \le 2\mu, \tag{5}$$

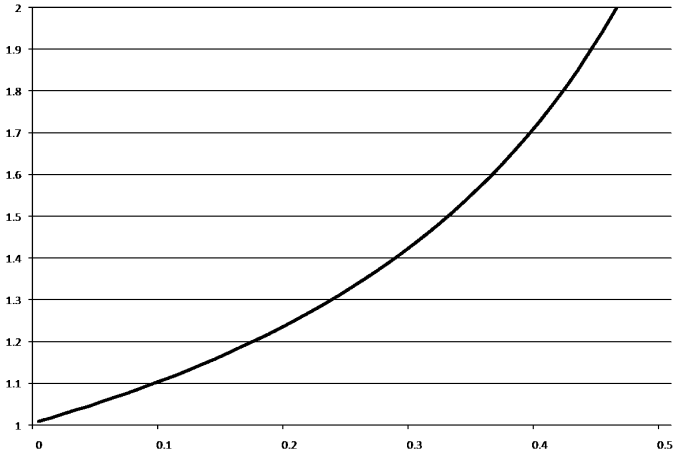


Figure 4. Ratio of μ_L over the mean of $F_1(t)$ (μ_1) for nanotube mean lengths up to 50% of grid size.

where μ is the mean nanotube length. In this case the mean for $F_1(t)$ is

$$\mu_1 = \mu\left(\frac{3\pi - 4\mu}{3\pi - 12\mu}\right).$$

Figure 4 shows the ratio mean for $F_L(t)$ over the mean of $F_1(t)$.

Once the mean nanotube length increases to 20% of the length of the grid, the increase in μ_1 will quickly get out of hand. We further investigate the effect of length biasing for the nanotube data in the following section, where we consider nonparametric estimation of the underlying length distribution.

3. ESTIMATING THE LENGTH DISTRIBUTION

Although the bias for $dF_1(t)$ in (4) is not the commonly observed form of length bias, it nonetheless presents a straightforward selection-bias estimation problem that can be solved directly. If our observed sample is denoted by $x_1, \ldots, x_n$, then, from Vardi (1985), the biased distribution can be expressed as some distribution function $G(t)$ such that

$$G(t) = \frac{\int_0^t w(u)\, dF_L(u)}{\int_0^\infty w(u)\, dF_L(u)},$$

where the weight function representing the bias is $w(x) = a\pi - 4\lambda x$. In this case $G(t)$ is identifiable only for values of $t < a\pi/(4\lambda)$.

As a result, it can be shown, from theorems 1 and 2 of Vardi (1985), that the NPMLE of F_L is

$$\hat{F}_L(t) = \frac{\sum_{i=1}^n I(x_i \le t)(a\pi - 4\lambda x_i)^{-1}}{\sum_{i=1}^n (a\pi - 4\lambda x_i)^{-1}}. \tag{6}$$

3.1 Properties of $\hat{F}_L$

Borrowing the method of proof from theorem 3.2 of Vardi (1982), we can show the following result. We assume the coefficient representing nanotube curvature λ is fixed and known.

Theorem 1. Assume that $F(t)$ is continuous for values $0 < t < a\pi/(4\lambda)$ and has bounded density with $F(0) = 0$ and $F(a\pi/(4\lambda)) = 1$. Then

$$\int_0^{a\pi/(4\lambda)} (a\pi - 4\lambda x)\, d\hat{F}_L(x) \longrightarrow \int_0^{a\pi/(4\lambda)} (a\pi - 4\lambda x)\, dF_L(x) \quad \text{with probability 1,}$$

and

$$\sqrt{n}(\hat{F}_L(t) - F_L(t))$$

converges in distribution to a pinned Gaussian process with mean 0 and covariance function $C(u, v)$, where, for $u \le v \le r$ and $\psi(t) = \int_0^t (a\pi - 4\lambda s)^{-1}\, dF_L(s)$,

$$C(u, v) = (a\pi - 4\lambda\mu_L) \times \big(\psi(u)(1 - F_L(v)) + F_L(u)(\psi(r)F_L(v) - \psi(v))\big).$$

3.2 Nanotube Data

In the analysis of the data of Wang et al. (2006), we use $\lambda = 1$ and refrain from making more elaborate assumptions about the apparent nanotube curvature displayed in the AFM image. Figure 5 shows the estimated survival function along with an approximate 95% confidence interval, based on the nonparametric percentile bootstrap procedure (with 1,000 bootstrap samples) described in chapter 15.2 of Kvam and Vidakovic (2007). Because of the large sample size, the coverage probability is not significantly improved using bias correction.

To see how the corrected nonparametric estimator compares with the uncorrected one, we can visually compare their density estimates. Figure 6 shows a smoothed version of the density associated with $\hat{F}_L(t)$ (dotted line) alongside the biased estimator constructed from the unadjusted empirical density (solid line). The estimated mean and standard deviation using the standard estimator (530.5 nm, 261.4 nm) are not remarkably different from the mean and standard deviation of the data (510.4 nm, 253.2 nm). In this case the mean nanotube length is on the order of 10 times smaller than the length of the side of the square from which they are sampled.

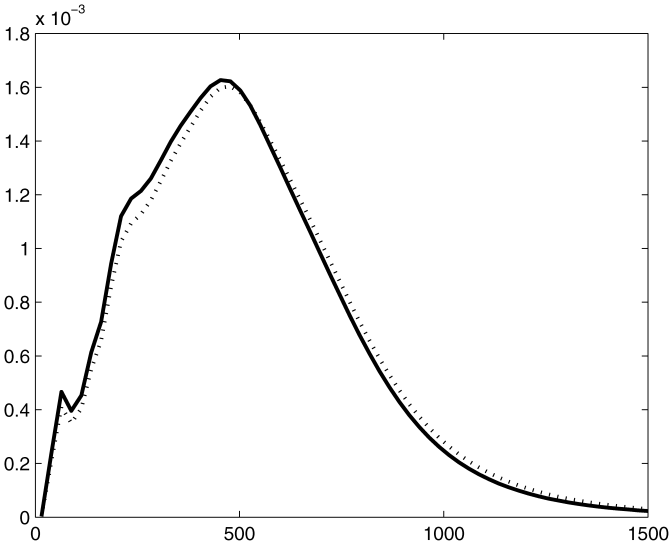


Figure 6. Empirical density (——) and adjusted density estimator (· · ·) for carbon nanotube lengths.

Obviously, if the nanotube lengths are more in proportion to the side of the square, then the length bias can be much more dramatic. Figure 7 shows the original empirical density and estimators adjusted for two smaller (hypothetical) AFM image squares. The squares have sizes measuring 2,500 and 2,000 nm (roughly between four and five times larger than the mean nanotube size). The estimated densities show how the NPMLE adjusts for censoring of more of the larger nanotubes as the window size is reduced. For the 2,500 nm^2 window, this is reflected in mass being transferred to the right, illustrating how substantial bias can be created by the missing censored observations.

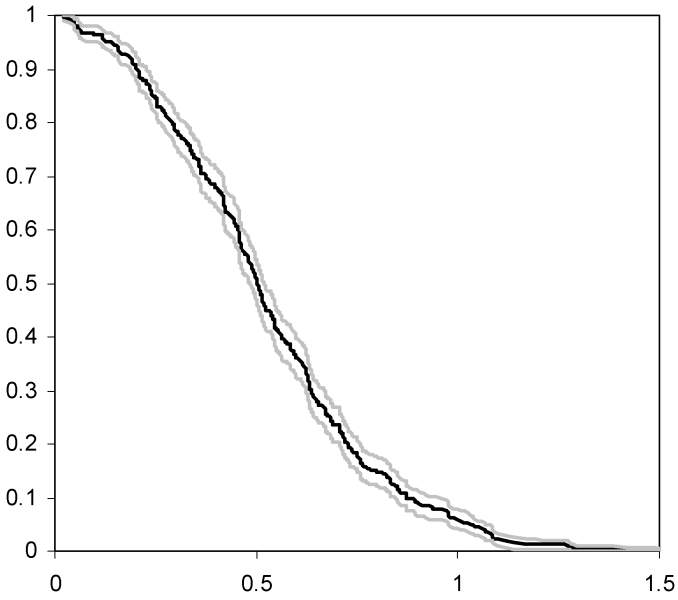


Figure 5. Nonparametric survival function (black line) and 95% confidence interval (gray lines) for carbon nanotube lengths.

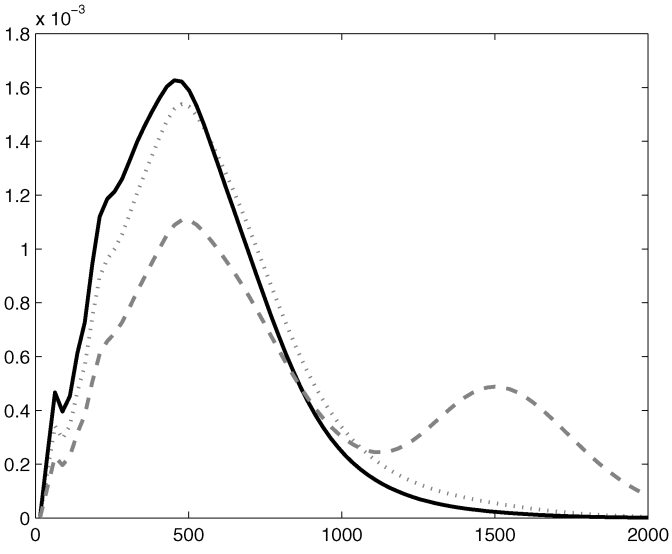


Figure 7. Empirical density (——) and adjusted density estimators based on different AFM image sizes. Original size is 5,000 nm. Adjusted sizes are 2,500 nm (· · ·) and 2,000 nm (- - -).

For the 2,000 nm^2, the NPMLE has generated an extra mode for probability mass representing the lost observations due to censoring.

3.3 Effect of Shape

With sound knowledge about the curvature of nanotubes in the AFM image, models based on $\lambda < 1$ suggest that the estimation of the length distribution can be improved. For the carbon nanotube data, however, changing the model by assigning $\lambda = .9$ will not change the NPMLE noticeably; for example, the estimated mean changes by only .68%. Estimating the sample curvature has not been addressed in the nanomanufacturing literature, even though it poses an intriguing problem for future research. More importance has been placed on tube lengths, diameters, and clusters than on how the nanotubes curve and bend.

Along with unknown curvature, heterogeneity within synthesized batches of carbon nanotubes has been carefully researched (see, e.g., Lu and Bhattacharya 2005). Again, the randomness of the nanotube curvature has not received special attention beyond how the tubes will overlay one another, which naturally affects the electrical properties of nanodevices. Although the metric implied in Figure 3 is overly simplistic in many regards, here it adequately serves the purpose of illustrating the effects of shape variability on the censoring, length-biasing, and nonparametric estimation of nanotube length.

In this example the carbon nanotubes are represented as bisected needles, and the angle of bend θ is governed by a uniform distribution, $\theta \sim \mathrm{U}(0, \pi/4)$. The resulting distribution of λ (related to the arc since distribution) percolates uncertainty down through $\hat{F}_L$. In this example the length distribution is $\mathrm{U}(0, 2\mu)$, and F_1 is represented as a mixture with mixing density $g_\lambda(w) = 4\pi^{-1}(1 - w^2)^{-1/2}$, corresponding to $\lambda = (\cos(\theta) + 1)/2$. Through numerical integration, we can compare the variation in noncensored observations from the mixture. If $\mu \leq .25$, then the standard deviation of the mixture sample is within 1% of the sample with fixed $\lambda = \pi/8$, suggesting that mild differences in nanotube curvature will not significantly affect the distribution of noncensored observations.

4. DISCUSSION OF RESULTS

The problem of locating and measuring nanosized matter has been greatly improved in recent years due to improving AFM technology and recognition software. But if the sampling mechanism is unable to include observations crossing the boundary of the observation grid, then the resulting length bias in the data set must be dealt with. If the bias is not addressed, then the length distribution for nanotubes will be underestimated (along with the variability). In the example of Wang et al. (2006), the consequences of the bias are relatively benign; however, other examinations of synthesized carbon nanotubes have included smaller regions relative to the average size of the nanotubes. The length of synthesized nanotubes are increasing as the technology improves; Burke, Li, and Yu (2006) describe carbon nanotubes of .4 cm in length that have been developed in research labs. Our analysis shows that the bias will have a substantial effect on the estimator of the length distribution if the mean nanotube length is as high as 20% of the size of the side of the square from which they are sampled.

Note that if censored observations were measured, then the analysis would be straightforward using the Kaplan–Meier product limit estimator. In some cases a nanotube located in the corner of a square can cross both grid boundaries, but in the nature of sampling nanotube lengths, the (right) censoring is the same.

Wang et al. (2006), set out to prove that the underlying length data can be characterized by the Weibull distribution. Indeed, the Weibull model has served as a precedent for similar distributional properties of fibrous materials. Based on the model for fiber composites of Fukuda and Kawata (1974) and Jayaraman and Kortschot (1996), Wang et al. (2006) showed that the epoxy composite stiffness measurements are consistent with Weibull-distributed nanotube lengths. Although their probability plot of the original data in Figure 8 is not entirely convincing, one may conclude that the Weibull presents a "reasonable" distribution fit to the sample. Knowing that the data are length-biased, we can reconsider this goodness of fit for the NPMLE, $\hat{F}_L$. Comparing the original Weibull probability plot in Figure 8 to the corrected one in Figure 9, we see that the Weibull fit actually has improved slightly. This is confirmed by the decrease in the computed Kolmogorov–Smirnov test statistics; that is, the corrected estimator for the length-biased data offers stronger evidence for the conjecture of Wang et al. (2006).

ACKNOWLEDGMENTS

This research was supported by National Science Foundation grant CMMI-0700131. The author thanks the associate editor and referees for their helpful comments that led to several improvements to the article.

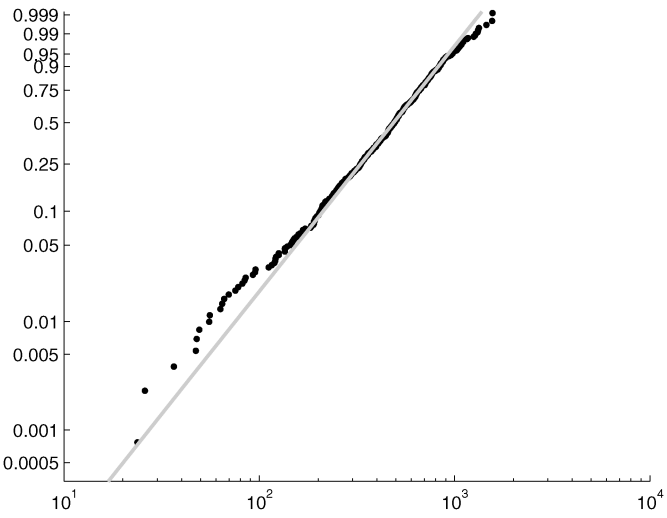


Figure 8. Original Weibull probability plot of the length-biased data for carbon nanotube lengths.

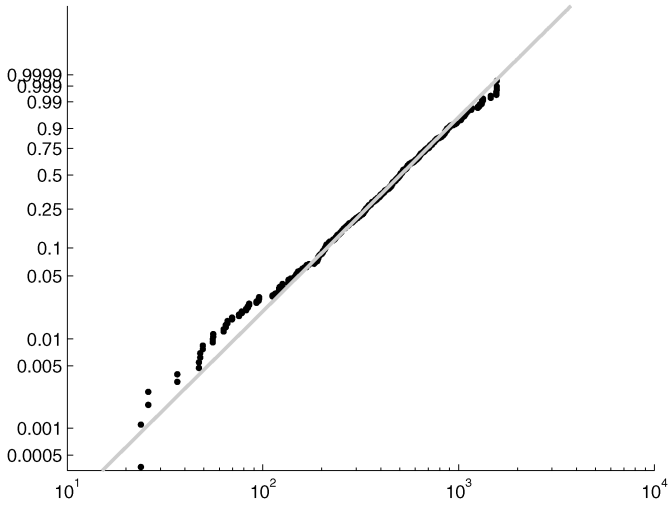


Figure 9. Weibull probability plot of the adjusted estimator of the carbon nanotube length distribution.

[Received August 2007. Revised July 2008.]

REFERENCES

Asgharian, A., M'Lan, C. E., and Wolfson, D. B. (2002), "Length-Biased Sampling With Right Censoring: An Unconditional Approach," *Journal of the American Statistical Association*, 97, 201–209.

Burke, P. J., Li, S., and Yu, J. (2006), "Quantitative Theory of Nanowire and Nanotube Antenna Performance," *IEEE Transactions on Nanotechnology*, 5, 314–334.

Fukuda, H., and Kawata, K. (1974), "On Young's Modulus of Short Fiber Composites," *Fibre Science and Technology*, 7, 207–212.

Gnedenko, B. V. (1962), *The Theory of Probability*, Moscow: MIS.

Horvath, L. (1985), "Estimation From a Length-Biased Distribution," *Statistics and Decisions*, 3, 91–113.

Jayaraman, K., and Kortschot, M. T. (1996), "Correction to the Fukuda–Kawata Young Modulus Theory and the Fukuda–Chao Strength Theory for Short Fiber–Reinforced Composite Materials," *Journal of Materials Science*, 31, 2059–2064.

Kvam, P. H., and Vidakovic, B. (2007), *Nonparametric Statistics With Applications to Science and Engineering*, Hoboken, NJ: Wiley.

Lu, Q., and Bhattacharya, B. (2005), "The Role of Atomistic Simulations in Probing the Small-Scale Aspects of Fracture: A Case Study on a Single-Walled Carbon Nanotube," *Engineering Fracture Mechanics*, 72, 2037–2071.

Mantel, N. (1952), "An Extension of the Buffon Needle Problem," *Annals of Mathematical Statistics*, 24, 674–677.

Perlman, M. D., and Wichura, M. J. (1975), "Sharpening Buffon's Needle," *The American Statistician*, 29, 157–163.

Ramaley, J. F. (1969), "Buffon's Noodle Problem," *American Mathematical Monthly*, 76, 916–918.

Rao, C. R. (1965), "On Discrete Distributions Arising Out of Methods of Ascertainment," in *Classical and Contagious Discrete Distributions*, ed. G. P. Patil, Calcutta, India: Pergamon Press and Statistical Publishing Society, pp.320–332.

Sun, J. (2006), *The Statistical Analysis of Interval-Censored Failure Time Data*, New York: Springer.

Vardi, Y. (1982), "Nonparametric Estimation in the Presence of Length Bias," *The Annals of Statistics*, 10, 616–620.

——— (1985), "Empirical Distributions in Selection Bias Models," *The Annals of Statistics*, 13, 178–203.

Wang, S., Liang, Z., Wang, B., and Zhang, C. (2006), "Statistical Characterization of Single-Wall Carbon Nanotube Length Distribution," *Nanotechnology*, 17, 634–639.

Ziegler, K. J., Gu, Z., Peng, H., Flor, E. L., Hauge, R. H., and Smalley, R. E. (2005), "Controlled Oxidative Cutting of Single-Walled Carbon SENTs," *Journal of the American Chemical Society*, 127, 1541–1547.